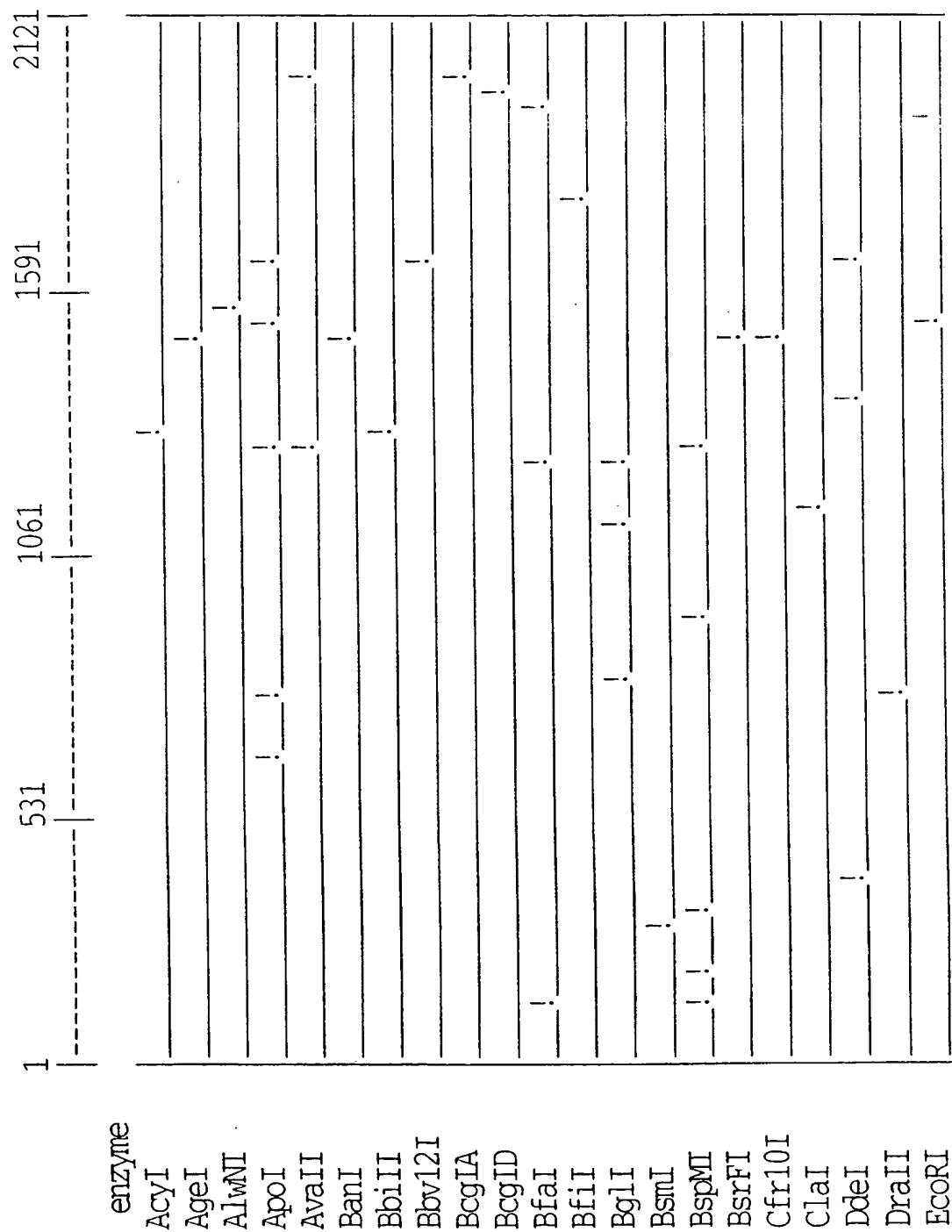


1/73

FIG. 1A

Restriction map of *M. catarrhalis* strain M35 *tbpB* gene

2/73

FIG. 1B

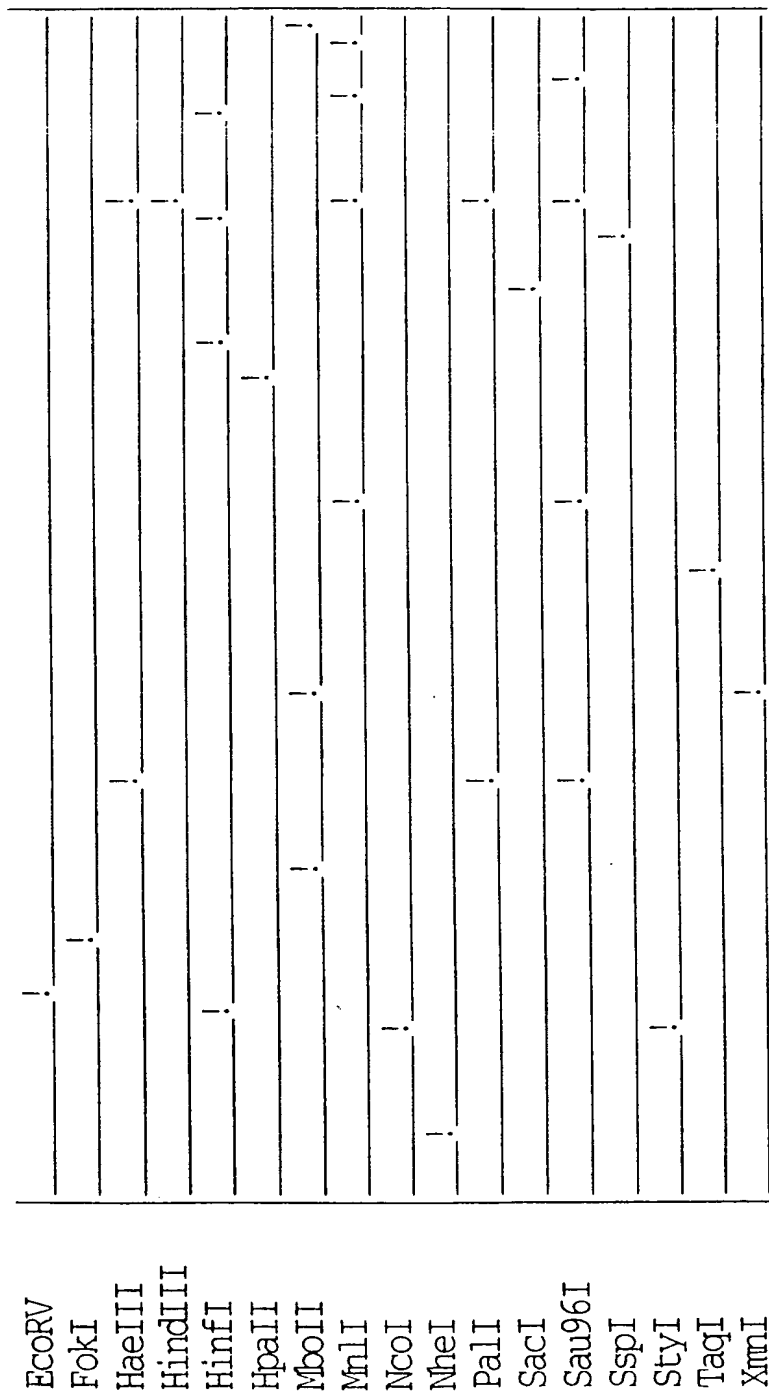


FIG.2A

M. catarrhalis strain M35 *tbpB* sequence

```

MET  LYS  HIS  ILE  PRO  LEU  THR  THR  LEU  CYS  ...
ATGAACACACTTCCTTTAACCACTGTGT...
10                                     20
...  VAL  ALA  ILE  SER  ALA  VAL  LEU  THR  ALA
... GTGGCAATCTCTGCCGTCCTTATTACCGCT
40                                     50
...                                     60

CYS  GLY  GLY  SER  GLY  GLY  SER  ASN  PRO  PRO  ...
TGTGGTGCGAGTGGTGGTTCTCAATCCACCT...
70                                     80
...  ALA  PRO  THR  PRO  ILE  PRO  ASN  ALA  SER  GLY
... GCTCCTACGCCCAATTCCAAATGCTAGCGGT
100                                    110
...                                     120

SER  GLY  ASN  THR  GLY  ASN  THR  GLY  ASN  ALA  ...
TCAGGTAAATACCTGGCAACACTGGTAATGCT...
130                                    140
...  GLY  GLY  THR  ASP  ASN  THR  ALA  ASN  ALA  GLY
... GCGGTACTGTATAATACAGCCAAATGCAAGT
150                                    160
...                                     170
...                                     180

ASN  THR  GLY  GLY  THR  ASN  SER  GLY  THR  GLY  ...
AATACAGCGGTACAAACTCTGGTACAGGC...
190                                    200
...  SER  ALA  ASN  THR  PRO  GLU  PRO  LYS  TYR  LYS
... AGTGCCAACACACCAAGAACCAATAATAA
210                                    220
...                                     230
...                                     240

```

FIG.2B

```

ASP VAL PRO THR ASP GLU ASN LYS LYS ASP ...
G A T G T C C A A C C G A T G A A A A T A A A A G A T ...
250
... GLU VAL SER GLY ILE GLN GLU PRO ALA MET
... G A A G T G T C A G G C A T T C A A G A A C C T G C C A T G
260
...
270...
280
...
290
300

GLY TYR GLY MET ALA LEU SER LYS MET ASN ...
G G T T A T G G C A T G G C T T T G A G T A A A A T G A A T ...
310
... LEU HIS LYS GLN GLN ASP THR PRO LEU ASP
... C T A C A C A A C A C A C A G A C A C G C C A T T A G A T
320
...
330...
340
...

GLU LYS ASP ILE ILE THR LEU ASP GLY LYS ...
G A A A G A T A T C A T T A C C T T A G A C G G T A A A ...
350
... LYS GLN VAL ALA LYS GLY GLU LYS SER PRO
... A A C A A G T T G C A A A A G G T G A A A A T C G C C A
360
...
370
380
390...
400
...

LEU PRO PHE SER LEU ASP VAL GLU ASN LYS ...
T T G C C A T T T C G T T G G A T G T A G A A A T A A A ...
410
420
430
440
450...
... LEU LEU ASP GLY TYR ILE ALA LYS MET ASN
... T T G C T T G A T G G C T A T A T A G C A A A A A T G A A T
460
470
480

```

FIG.2C

GLU ALA ASP LYS ASN ALA ILE GLY ASP ARG ...
 G A G C G G A T A A A A T G C C A T T G G T G A C A G A ...
 490 500 510...
 ... ILE LYS LYS ASP ASN LYS ASP LYS SER LEU
 ... A T T A A G A A A G A T A A T A A G A C A A G T C A T T A
 520 530 540
 ...

SER LYS ALA GLU LEU ALA LYS GLN ILE LYS ...
 T C T A A A G C A G A G C T T G C C A A A C A A A T C A A A ...
 550 560 570...
 ... GLU ASP VAL ARG LYS SER HIS GLU PHE GLN
 ... G A G A T G T G C G T A A A A G C C A T G A G T T C A G
 580 590 600
 ...

5/73

GLN VAL LEU SER SER LEU LYS ASN LYS ILE ...
 C A G T A T T A T C A T C A C T G A A A A C A A A T T ...
 610 620 630...
 ... PHE HIS SER ASN ASP GLY THR THR LYS ALA
 ... T T T C A T T C A A A T G A T G G A A C A C C A A A G C A
 640 650 660
 ...

THR THR ARG ASP LEU GLN TYR VAL ASP TYR
 A C C A C G A G A T T T A C A A T A T G T T G A T T A T
 670 680 690
 ... GLY TYR TYR LEU VAL ASN ASP GLY ASN TYR
 ... G G T T A C T A C T T G G T G A A T G A T G G C A A T T A T
 700 710 720
 ...

FIG.2D

```

LEU THR VAL LYS THR ASP GLU LEU TRP ASN ...
CT : ACCGTCAAACAGACGAACTTTGGAAAT
730
... LEU GLY PRO VAL GLY GLY VAL PHE TYR ASN
... TTAGGCCCTGTGGCGGGTGTGTTTATAAT
760
...
770
...
780

GLY THR THR THR ALA LYS GLU LEU PRO THR ...
GGCAACAACGACCGCCAAAGAGCTACCCACA...
790
... GLN ASP ALA VAL LYS TYR LYS GLY HIS TRP 6/73
... CAGATGCGGTCAAATAATAAGGACATTCG
800
...
820
...

ASP PHE MET THR ASP VAL ALA LYS GLN ARG ...
GACTTTATGACCGATGTTGCCCAACAAGAA...
850
... ASN ARG PHE SER GLU VAL LYS GLU ASN LEU
... AACCGATTTAGCGAAGTGAAAGAAACCTT
860
...
880
...

GLN ALA GLY ARG TYR TYR GLY ALA SER SER ...
CA : GCAGGTCGGGTATTATGGAGCATCTCA...
910
... LYS ASP GLU TYR ASN ARG LEU LEU THR ASP
... AAGATGAATACACCGCTTATACTGAT
920
...
940
...
950
...
960

```

FIG.2E

GLU LYS ASN LYS PRO GLU ARG TYR ASN GLY ...
 G A G A A A A C A A C C A G A G C G T T A T A C G G T ...
 970 980 990...
 ... GLU TYR GLY HIS SER SER GLU PHE THR VAL
 ... G A A T A T G G T C A T A G C A G T G A G T T T A C T G T T
 1000 1010 1020
 ...

ASN PHE LYS LYS LYS LYS LEU THR GLY GLU ...
 A A T T T A A G G A C A A A A A A T T A C A G G T G A G ...
 1030 1040 1050
 ... LEU PHE SER ASN LEU GLN ASP SER ARG LYS
 ... C T G T T A G T A A C C T A C A A G A C A G C C G T A A G
 1060 1070 1080
 ...

7/73

GLY ASN VAL THR LYS LYS THR LYS ARG TYR ASP ...
 G G C A A T G T T A C G A A A C C A A A C G C T A T G A C ...
 1090 1100 1110...
 ... ILE ASP ALA ASN ILE TYR GLY ASN ARG PHE
 ... A T C G A T G C C A A T A T C T A C G G C A A C C G C T T C
 1120 1130 1140
 ...

ARG GLY SER ALA THR ALA SER ASP LYS ALA ...
 C G T G G C A G T G C C A C C G C A A G C G A T A A G C A ...
 1150 1160 1170...
 ... GLU ALA SER LYS THR LYS HIS PRO PHE THR
 ... G A G C A A G C A A A C C A A A C A C C C C T T T A C C
 1180 1190 1200
 ...

FIG.2F

```

SER  ASP  ALA  LYS  ASN  SER  LEU  GLU  GLY  GLY  ...
AGCGATGCCAAATAAGCCTAGAGGCGGT...
1210
...  PHE  TYR  GLY  PRO  ASN  ALA  GLU  LEU  ALA
...  TTTTATGGGACCAACGCCCGAGGAGCTGGCA
...                                     1250
...                                     1260

GLY  LYS  PHE  LEU  THR  ASN  ASP  ASN  LYS  LEU  ...
GGTAATAATTCCTAACCAATGACACAACCTC...
1270
...  PHE  GLY  VAL  PHE  GLY  ALA  LYS  ARG  GLU  SER
...  TTTGGCGTCTTTGGTGCTAAACGAGAGAGT
...                                     1300
...                                     1310
...                                     1320

LYS  ALA  GLY  GLU  LYS  THR  GLU  ALA  ILE  LEU  ...
AAGCTGGGGAATAAACCGAAGCCATCTTA...
1330
...  ASP  ALA  TYR  ALA  LEU  GLY  THR  PHE  ASN  LYS
...  GATGCCCTATGCACCTTGGGACATTTAACAAA
...                                     1360
...                                     1370
...                                     1380

ASN  ASN  ALA  THR  THR  PHE  THR  PRO  PHE  THR  ...
AATACGCAACCAACATTCACCCCATTTACC...
1390
...  LYS  LYS  GLN  LEU  ASP  ASN  PHE  GLY  ASN  ALA
...  AAAACAACATGGATATACTTTGGCAATGCC
...                                     1420
...                                     1430
...                                     1440

```


FIG.2G

```

LYS  LYS  LEU  VAL  LEU  GLY  SER  THR  VAL  ILE ...
A A A A G T T G G T C T T G G G T T C T A C C G T C A T T ...
1450
... ASP  LEU  VAL  PRO  THR  GLY  VAL  THR  LYS  ASP
... G A T T G G T G C C T A C C G G T G T C A C C A A A G A T
1460
...
1470
...
1480
...
1490
1500

VAL  ASN  GLU  PHE  THR  LYS  ASN  LYS  PRO  ASP ...
G T C A A T G A A T T C A C C A A A A C A A G C C A G A T ...
1510
... SER  ALA  THR  ASN  LYS  ALA  GLY  GLU  THR  LEU
... T C T G C C A C A A C A A A G C G G C G A G A C T T T G
1520
...
1530
...
1540
1550
1560
9 / 73

MET  VAL  ASN  ASP  LYS  VAL  SER  VAL  LYS  THR ...
A T G G T G A A T G A T A A A G T T A G C G T C A A A C C ...
1570
... TYR  GLY  TYR  GLY  ARG  ASN  PHE  GLU  TYR  LEU
... T A T G G C T A T G G C A G A A A C T T T G A A T A C C T A
1580
...
1590
1600
1610
1620

LYS  PHE  GLY  GLU  LEU  SER  VAL  GLY  THR  SER ...
A A A T T G G T G A G C T C A G T G T C G G C A C A G C ...
1630
... ASN  SER  VAL  PHE  LEU  GLN  GLY  GLU  ARG  THR
... A A C A G C G T C T T T T A C A A G G C G A A C G C A C C
1640
...
1650
1660
1670
1680

```

FIG.2H

ALA THR THR GLY GLU LYS ALA VAL PRO THR ...
 GCTACCAAGGCGAGAAAGCCGTACCAACC...
 1690 1700 1710...
 ... LYS GLY THR ALA LYS TYR LEU GLY ASN TRP
 ... AAGGCACAGCCCAAATAATTGGGGAAC TTGG
 1720 1730 1740
 ...

VAL GLY TYR ILE THR GLY LYS ASP SER ...
 GTAGGATATCATCATCAGGAAGGACTCATCA...
 1750 1760 1770...
 ... LYS SER PHE ASN GLU ALA GLN ASP VAL ALA
 ... AAGCTTTAATGAGGCCCAAGATGTTGCT
 1780 1790 1800
 ...

10/73

ASP PHE ASP ILE ASP PHE GLU LYS LYS SER ...
 GATTTGACATTGACTTTTGAGAAATAATCA...
 1810 1820 1830...
 ... VAL LYS GLY LYS LEU THR THR LYS ASP ARG
 ... GTTAAAGGCATAACTGTACCAACCACCAAGACCGC
 1840 1850 1860
 ...

GLN ASP PRO VAL PHE ASN ILE THR GLY ASP ...
 CAGACCCCTGTATTTAACATCAGGTGAC...
 1870 1880 1890...
 ... ILE ALA GLY ASN GLY TRP THR GLY LYS ALA
 ... ATCGCAGGCATAATGGCTGGACACGCAAGCC
 1900 1910 1920
 ...

FIG.2I

```

SER  THR  THR  LYS  ALA  ASP  ALA  GLY  GLY  TYR  ...
AGC ACC ACC A A A G C G G A C G C A G G G G C T A C ...
1930
...  LYS  ILE  ASP  SER  SER  THR  GLY  LYS  SER
...  A A G A T A G A T T C T A G C A G T A C A G G C A A A T C C
1940
...  1960
...  1970
...  1980

ILE  VAL  ILE  LYS  ASP  ALA  GLU  VAL  THR  GLY  ...
ATC GTC ATCA A A G A T G C C C G A G G T T A C A G G G ...
1990
...  GLY  PHE  TYR  GLY  PRO  ASN  ALA  ASN  GLU  MET
...  GGC TTT TAT GGT CCA AAT GCA A A C G A G A T G
2000
...  2020
...  2030
...  2040

GLY  GLY  SER  PHE  THR  HIS  ASN  THR  ASP  ASP  ...
GGC GGT CAT TTA CACA CACA CACCGATGAC...
2050
...  SER  LYS  ALA  SER  VAL  VAL  PHE  GLY  THR  LYS
...  AGTAAAGCCCTCTGTGGTCTTTGGCACA A A A
2060
...  2070
...  2080
...  2090
...  2100

ARG  GLN  GLU  GLU  VAL  LYS  ***
AGACAAGAGAGTTAGTAG
2110
2120

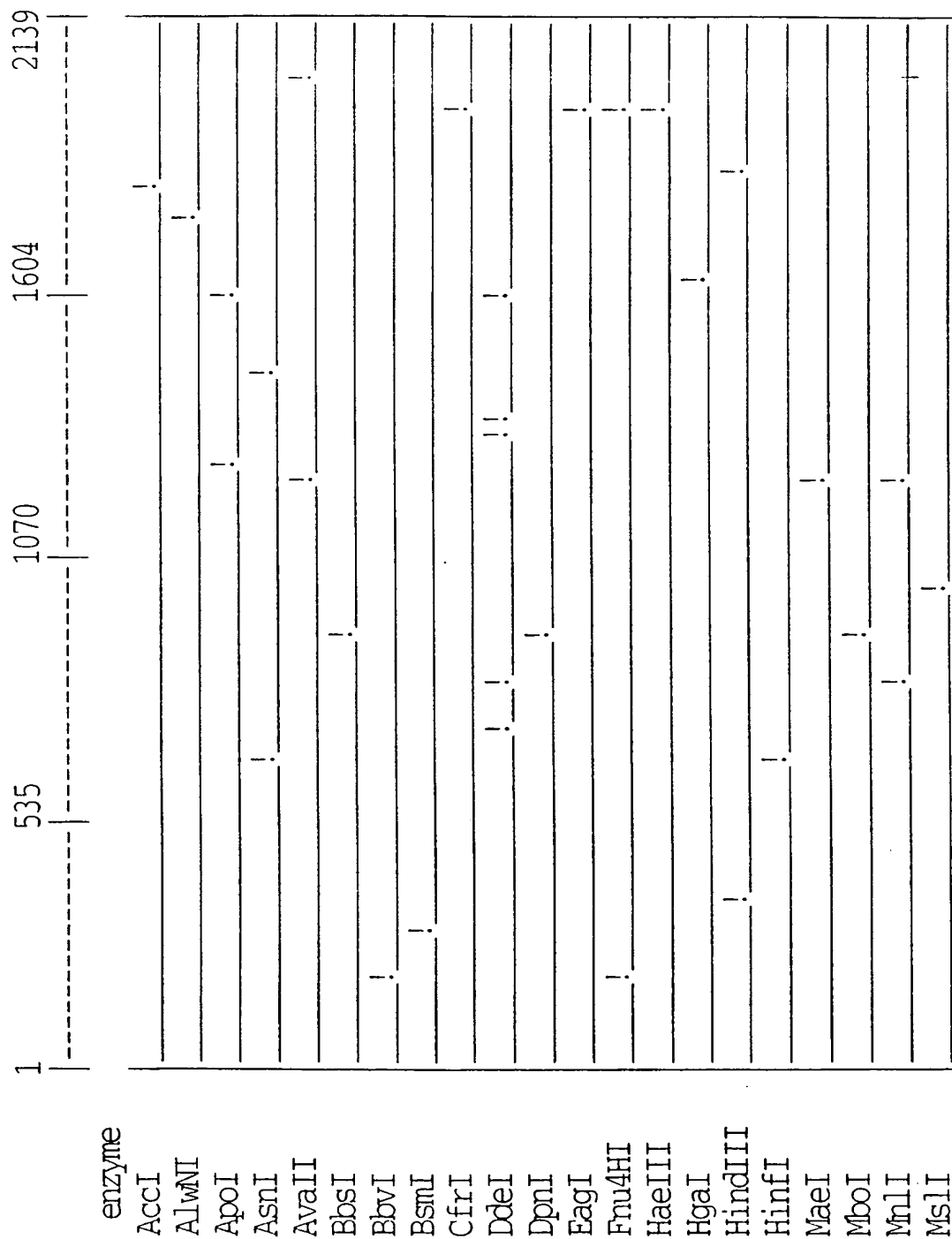
```

11/73

12/73

FIG.3A

Restriction map of *M. catarrhalis* strain 3 *tbpB* gene



[illegible]

FIG. 3B

FIG.4A

M. catarrhalis strain 3 *tbpB* sequence

```

MET  LYS  HIS  ILE  PRO  LEU  THR  THR  LEU  CYS  ...
ATGAACACATTCTTAAACCACACTGTGT...
10                                     20
... VAL  ALA  ILE  SER  ALA  VAL  LEU  LEU  THR  ALA
... GTGGCAATCTCTGCCGCTCTATTAAACCGCT
40                                     50
...                                     60

```

```

CYS  GLY  GLY  SER  GLY  GLY  SER  ASN  PRO  PRO  ...
TGTGGTGGCAGTGGTGTTCATAATCCACT...
70                                     80
... ALA  PRO  THR  PRO  ILE  PRO  ASN  ALA  GLY  GLY
... GCTCCTACGCCCATTCCTCAATGCAGGCGGT
100                                    110
...                                     120

```

```

ALA  GLY  ASN  ALA  GLY  SER  GLY  THR  GLY  GLY  ...
GAGGTAAATGCTGGTAGCGGTACTGGCGGT...
130                                    140
... ALA  GLY  SER  THR  ASP  ASN  ALA  ALA  ASN  ALA
... GCAGGTAGCACTGATAATGCCAATGCA
160                                    170
...                                     180

```

```

GLY  SER  THR  GLY  GLY  ALA  SER  SER  GLY  THR  ...
GGCAGTACAGCGGTGCAAGCTCTGGGTACA...
190                                    200
... GLY  SER  ALA  SER  THR  GLN  LYS  PRO  LYS  TYR
... GGCAGTGCCAGCACACAAATAAATAAT
220                                    230
...                                     240

```

SUBSTITUTE SHEET (RULE 26)

09/673133

FIG.4C

```

GLU  LYS  GLN  ASN  ILE  GLU  ASN  GLN  ILE  LYS  ...
G A G A G C A A A C A T T G A A A A T C A A A T C A A A ...
490
...  LYS  GLU  ASN  LYS  GLU  LEU  ASP  LYS  THR  ALA
...  A A G A A A A T A A A G A A C T T G A T A A A C G G C A
500
...
520
530
540

LEU  LYS  ALA  LEU  ILE  GLU  LYS  VAL  LEU  ASP  ...
C T A A A G C T C T T A T T G A A A A A G T T C T T G A T ...
550
...  ASP  TYR  LEU  THR  SER  LEU  ALA  LYS  PRO  ILE
...  G A C T A T C T A C A A G T C T T G C T A A A C C C A T T
560
...
580
590
600
610/73

TYR  GLU  LYS  ASN  ILE  ASN  ASP  SER  HIS  ASP  ...
T A T G A A A A A A A T A T T A A T G A T T C A C A T G A T ...
610
...  LYS  GLN  ASN  LYS  ALA  ARG  THR  ARG  ASP  LEU
...  A A G C A G A A T A A A G C A C G C A C T C G T G A T T G
620
...
640
650
660

LYS  TYR  VAL  ARG  SER  GLY  TYR  ILE  TYR  ARG  ...
A A G T A T G T G C G T T C T G T T A T T A T T A T C G C ...
670
...  SER  GLY  TYR  SER  ASN  ILE  ASP  ILE  GLN  LYS
...  T C A G G T T A T T C T A A T A T C G A C A T T C A A A G
680
...
700
710
720

```


FIG.4D

```

LYS ILE ALA LYS THR GLY PHE ASP GLY ALA ...
A A A T A G C C T A A A A C T G G T T T T G A T G G T G C T ...
730
... LEU PHE TYR LYS GLY THR GLN THR ALA LYS
... T T A T T T A T A A G G T A C A C A A A C T G C T A A A
760
...
770
780

GLN LEU PRO VAL SER GLU VAL LYS TYR LYS ...
C A A T G C C C T G T A T C T G A G G T T A G T A T A A A ...
790
... GLY THR TRP ASP PHE MET THR ASP ALA LYS
... G G C A C T T G G G A T T T A T G A C C G A T G C C A A A
800
...
810
820
830
840
17/73

LYS GLY GLN SER PHE SER SER PHE GLU ARG ...
A A G G A C A A T C A T T T A G C A G T T T T G A A A G A ...
850
... ARG ALA GLY ASP ARG TYR SER ALA MET SER
... C G A G C T G G T G A T C G C T A T A G T G C A A T G T C T
860
...
870
880
890
900

SER HIS GLU TYR PRO SER SER LEU LEU THR ASP ...
T C C C A T G A G T A C C C A T C T T T A T T A C T G A T ...
910
... ASP LYS ASN LYS PRO ASP ASN TYR ASN ASP
... G A T A A A A C A A C C A G A T A A T A T A C G A T
920
...
930
940
950
960

```

FIG.4E

GLU TYR GLY HIS SER SER GLU PHE THR VAL ...
 G A A T A T G G T C A T A G C A G T G A G T T T A C G G T A ...
 970 980 990...
 ... ASP PHE SER LYS LYS SER LEU THR GLY GLY
 ... G A T T T A G T A A A A G A G C C C T A C A G G T G G G
 1000 1010 1020
 ...

LEU PHE SER ASN LEU GIN ASP HIS HIS LYS ...
 C T G T T A G T A C C T A C A A G A C C A C C A T A A G ...
 1030 1040 1050...
 ... GLY LYS VAL THR LYS THR LYS ARG TYR ASP
 ... G G C A A G G T T A C G A A A C C C A A C G C T A T G A C
 1060 1070 1080 18/73
 ...

ILE ASN ALA ARG ILE HIS GLY ASN ARG PHE ...
 A T C A A T G C C C G T A T C C A C G G T A C C G C T T C ...
 1090 1100 1110...
 ... ARG GLY SER ALA THR ALA ILE ASN LYS ASP
 ... C G T G G C A G T G C C A C C G C A A T C A A T A A G A T
 1120 1130 1140
 ...

ASN GLU SER LYS ALA LYS HIS PRO PHE THR ...
 A A T G A A A G C A A A G C C A A A C A C C C C T T A C C ...
 1150 1160 1170...
 ... SER ASP ALA ASP ASN ARG LEU GLU GLY GLY
 ... A G C G A T G C C G A C A A T A G G C T A G A G G C G G T
 1180 1190 1200
 ...

FIG.4F

```

PHE TYR GLY PRO ASN ALA GLU LEU ALA ...
T T T A T G G A C C A A C G C C G A G G A G C T G G C A A ...
1210
... GLY LYS PHE LEU THR ASP ASP ASN LYS LEU
... G G T A A A T T C C C T A C C G A T G A C A A C A A C T C
1240
...
1250
1260

PHE GLY VAL PHE GLY ALA LYS GLN GLU SER ...
T T T G G T G T C T T T G G T G C T A A C A A G A G A G T ...
1270
... GLU ALA LYS GLU THR GLU ALA ILE LEU ASP
... G A A G C T A A G G A A C C G A A G C C A T C T T A G A T
1300
...
1310
1320

ALA TYR ALA LEU GLY THR PHE ASN LYS SER ...
G C T T A T G C A C T T G G G A C A T T A A T A A A I C T ...
1330
... GLY THR THR ASN PRO ALA PHE THR ALA ASN
... G G T A C G A C C A A T C C T G C C T T T A C C G C C A A T
1360
...
1370
1380

SER LYS LYS GLY LEU ASP ASN PHE GLY ASN ...
A G T A A A A G A A C T G G A T A A C T T T G G C A A T ...
1390
... ILE ASN LYS LEU VAL LEU GLY SER THR VAL
... A T T A A T A A A T T G G T C T T G G G T T C T A C T G T G
1420
...
1430
1440

```

FIG.4G

```

ILE ASP LEU THR GLN GLY ASN ASP PHE VAL ...
A TAG ACC T T A C T C A A G G T A A T G A T T T G T A ...
1450                                     1460 1470...
... LYS THR ILE ASP LYS GLU LYS PRO ALA THR
... A A A C C A T T G A T A A G A A A G C C A G C C A C C
1480                                     1490 1500
...

THR THR ASN GLN ALA GLY GLU PRO LEU THR ...
A C T A C C A A T C A A G C A G G C G A G C C T T T G A C G ...
1510                                     1520 1530...
... VAL ASN ASP LYS VAL ARG VAL GLN VAL CYS
... G T G A A T G A T A A G G T T C G G G T A C A G T T T G T
1540                                     1550 1560
...

CYS SER ASN LEU GLU HIS LEU LYS PHE GLY ...
T G T A G C A A T C T T G A G C A T C T A A A T T T G G C ...
1570                                     1580 1590...
... SER LEU SER ILE GLY ASP SER ASN SER VAL
... T C A C T G A G T A T C G G T G A T A G T A A T A G C G T C
1600                                     1610 1620
...

PHE LEU GLN GLY GLU ARG THR ALA THR LYS ...
T T T T A C A A G G T G A A C G C A C C G C T A C C A A A ...
1630                                     1640 1650...
... GLY ASP LYS ASP LYS ALA MET PRO VAL ALA
... G G T G A T A A A G A T A A A G C C A T G C C A G T T G C A
1660                                     1670 1680
...

```

FIG. 4H

GLY ASN ALA LYS TYR ARG GLY THR TRP ALA ...
 GGAATAATGCTAATAATACCGTGCTACATGGGCA...
 1690 1700 1710...
 ... GLY TYR VAL ALA GLY SER GLY ASN THR SER
 ... GGCTATGTGTGCAAGGCTCTGGCAATAACCAAGC
 1720 1730 1740
 ...

LYS ALA TYR GLU ALA GLN GLN PHE ALA ASP ...
 AAGCCTATGAAGCCCAACAATTGCTGAC...
 1750 1760 1770...
 ... ASN ALA ASN ARG ALA GLU PHE ASP VAL ASP
 ... AATGCCAACCGTGCCGAGTTTGATGTAGAC
 1780 1790 1800
 ...

21/73

PHE ALA ASN LYS SER LEU THR GLY LYS LEU ...
 TTGCTAACAAGCCCTAAGCTGCTT...
 1810 1820 1830...
 ... ILE PRO ASN THR SER SER ASP GLY LYS SER
 ... ATCCAAATAACGAGCAGTGATGATAATCT
 1840 1850 1860
 ...

ALA PHE ASP ILE THR ALA THR ILE ASP GLY ...
 GCTTTGATATTACTGCTACATAATGATGGC...
 1870 1880 1890...
 ... ASN GLY PHE SER GLY LYS ALA ASN THR PRO
 ... AATGGTTTATGTTGTAAGCCAATAACCA
 1900 1910 1920
 ...

FIG.4I

```

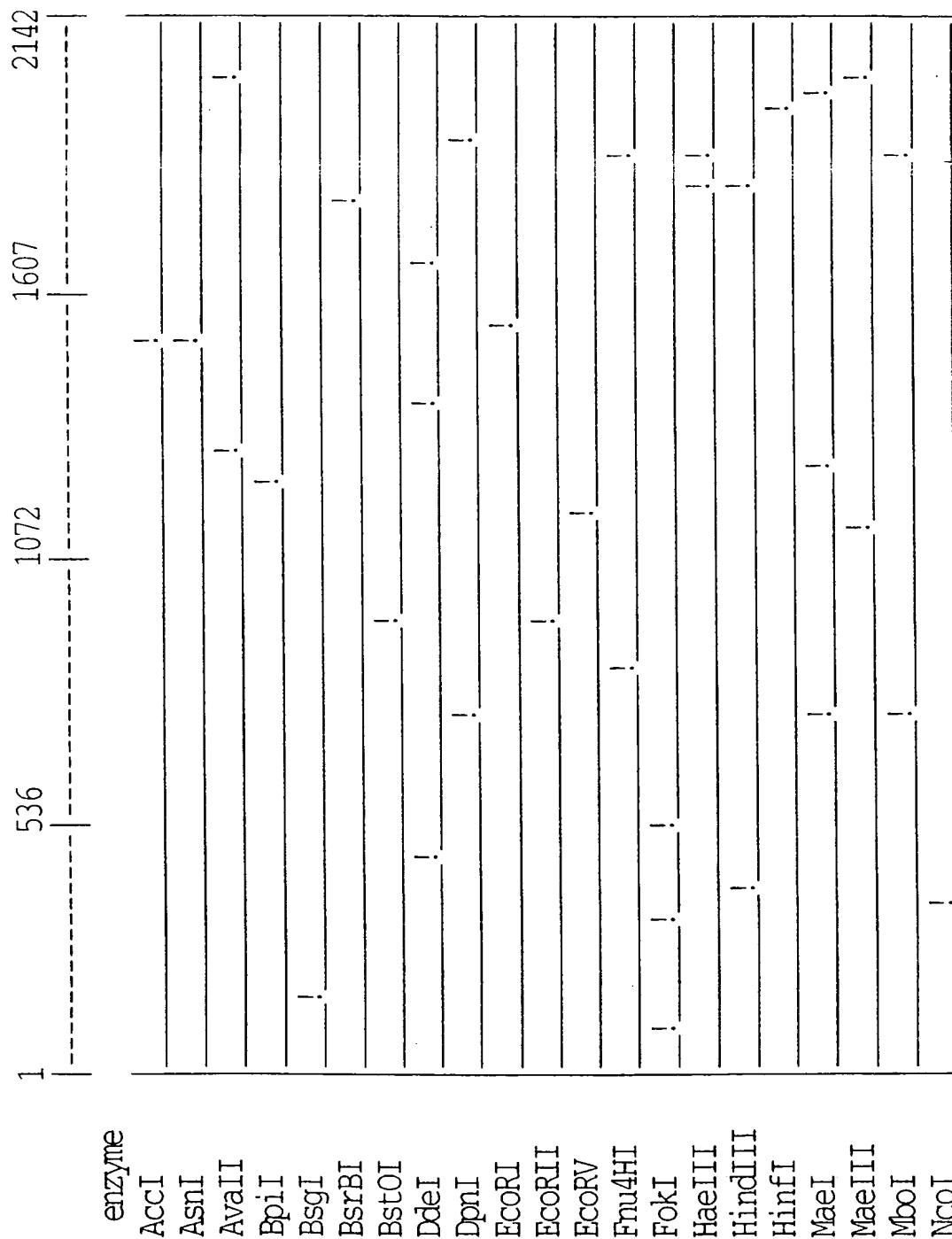
ASP  ILE  GLU  THR  GLY  GLY  LEU  LYS  ILE  ASP  ...
G A T A T T G A A C A G G T G G G T T A A G A T T G A C ...
1930
...  SER  LYS  ASN  SER  GLU  SER  GLY  ARG  VAL  ILE
...  A G T A A G A A C A G T G A A A G C G G C C G A G T A A T T
1940
...  1960
...  1970
...  1980
...  1990
...  2000
...  TYR  GLY  PRO  GLN  ALA  ASN  GLU  LEU  GLY  GLY
...  T A T G G T C C C A C A G C T A A T G A A C T G G G T G G C
...  2010
...  2020
...  2030
...  2040
...  2050
...  ASN  GLN  ASP  LYS  ASP  SER  SER  ALA  SER  VAL
...  A A T C A A G A C A A A G A C A G T A G T G C A T C T G T G
...  2060
...  2070
...  2080
...  2090
...  2100
...  2110
...  2120
...  2130
...  LYS  PRO  ***
...  A A A C C A T G A

```

23/73

FIG.5A

Restriction map of *M. catarrhalis* strain LES1 *tbpB* gene



24/73

FIG. 5B

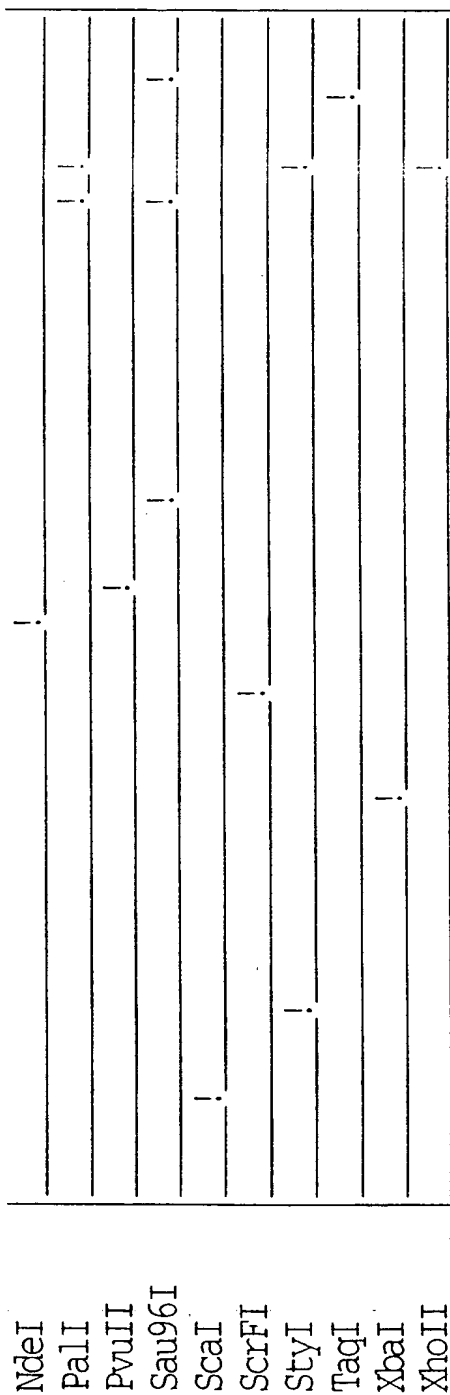


FIG.6A

M. catarrhalis strain LES1 *tbpB* sequence

```

MET  LYS  HIS  ILE  PRO  LEU  THR  THR  LEU  CYS  ...
ATGAACACACATTCTTAAACCACTGTGT...
10
... VAL  ALA  ILE  SER  ALA  VAL  LEU  LEU  THR  ALA
... GTGGCAATCTCTGCCGTCCTTATTAAACCGCT
20
...
40
60

```

```

CYS  GLY  GLY  SER  GLY  GLY  SER  ASN  PRO  PRO  ...
TGTGGTGGCAGTGCTGGTTCAATAATCCACT...
70
... ALA  PRO  THR  PRO  ILE  PRO  ASN  ALA  GLY  SER
... GCTCCTACGCCCATCCCAATGCAGGCAGT
80
...
100
110
120

```

```

ALA  GLY  ASN  ALA  GLY  GLY  THR  GLY  ASN  THR  ...
GCAGGTAAATGCTGGCGGTACAGGAATAACA...
130
... GLY  GLY  THR  GLY  SER  THR  ASP  ASN  VAL  GLY
... GCGGTACTGGCAGTACTGATATGTAGGC
140
...
160
170
180

```

```

ASN  ALA  GLY  GLY  ALA  ASN  SER  GLY  THR  GLY  ...
AATGCTGGCGGTGCCAAACTCTGTGTACAGGC...
190
... ASN  ALA  GLY  ASN  SER  GLY  ASN  ALA  ASN  SER
... AATGCAGGTAAATTCAGGTAAATGCAACTCT
200
...
220
230
240

```

FIG.6B

GLY THR GLY SER ALA ASN THR PRO GLU PRO ...
 GGTACAGGCAGTGCCACAACACCAAGAACCA...
 250 260 270...
 ... LYS TYR GLN ASP VAL PRO THR ASP LYS ASN
 ... AATAATCAAGATGTGCCCAACCGATAAATAAT
 ... 280 290 300

GLU LYS GLU GLN VAL SER SER ILE GLN GLU ...
 GAAAGAGACAGTTTCATCCATTCAAGAA...
 310 320 330...
 ... PRO ALA MET GLY TYR ALA MET GLU LEU LYS
 ... CCTGCCATGGGTTATGCAATGGAAATTAAAG
 ... 340 350 360

26/73

LEU ARG ASN ALA HIS PRO LEU ASN PRO ASN ...
 CTCGTAAATGCTCACCCCTCTTAACCCAAAT...
 370 380 390...
 ... LYS ASN LYS GLU ALA GLU LYS ARG ILE ALA
 ... AATAATAAGAGGCTGAATAACGCATTGCC
 ... 400 410 420

LEU ASP GLN LYS ASP LEU VAL ALA VAL GLU ...
 TTAGACCACAATAAGATTGTGGTGCAGTAGAG...
 430 440 450...
 ... GLY ASP LEU THR ASN ILE PRO PHE ASP LYS
 ... GGCGACCTAACCAACATTCTTTGATAAA
 ... 460 470 480

FIG.6C

```

ASN LEU ILE GLU TYR LEU LYS LYS SER SER ...
AATCTTATTGAAATACCTTTAAATAATCATCC...
490 500 510...
... GLU VAL VAL SER LYS PHE GLU ALA GLN LYS
... GAGGTGTAGTAATAATTGAGCAACAATAA 540
... 520 530 540

GLY GLY ILE GLU ASN ASN THR ARG LEU THR ...
GGCGGTATTGAAATAACACACAGACTGACA...
550 560 570...
... HIS LYS ASP LEU SER SER GLU GLN LYS GLU
... CACAAGATTTTATCATCATCAGAGCAATAAGAA 600
... 580 590 600

ALA LYS VAL LYS GLU ALA LEU ASP ASN ALA ...
GCAAGTCAAGAGAGCGTTGGACAAATGCT...
610 620 630...
... LEU THR GLN PHE ALA GLN GLU LYS TYR LYS
... TTAACCTCAATTGCCCCAAGAAATAACAAG 660
... 640 650 660

GLU LEU ILE GLU ASN ALA HIS ASP LYS LYS ...
GAGCTAATTGAGAACGCCCATGATAAATAA...
670 680 690...
... SER ASP ALA ARG ASN ARG ASP LEU GLU TYR
... TCTGACGCACGCACACCGTGATCTAGATAT 720
... 700 710 720

```

FIG.6D

```

VAL  LYS  SER  GLY  PHE  ASN  TYR  LEU  SER  GLY  ...
GTC  AAG  CTC  TGG  TTT  AAC  TAT  CTT  CTC  TGA  ...
730
...  TYR  THR  ALA  THR  ASP  HIS  ASP  LYS  LYS  THR
...  TAT  ACC  GCC  ACC  GAC  CAC  CAG  CAA  AAA  AAC  C
760
...
ASN  TYR  ARG  GLY  TYR  TYR  GLY  ALA  LEU  TYR...
AAT  ATC  GTG  GCT  ATAT  ATG  GTG  CGT  TGT  AT...
790
...  TYR  LYS  GLY  SER  GLU  THR  ALA  LYS  GLU  LEU
...  TAT  AAA  GGC  AGC  GAA  ACC  GCC  CAA  AGA  GCT  A
820
...
PRO  GLN  THR  SER  ALA  LYS  TYR  LYS  GLY  TYR  ...
CCA  CAA  ACA  AGT  GCA  AAA  ATA  TAA  AGT  TAT  ...
850
...  TRP  ASP  PHE  MET  THR  ASP  ALA  THR  LEU  ASP
...  TGG  GACT  TTT  ATG  ACA  GAT  GCC  ACA  CTT  GAT
880
...
ASN  LYS  TYR  THR  ASP  LEU  PRO  GLY  ILE  ALA  ...
ACA  AAA  ATA  CAC  GGA  TTT  GCC  CAG  GTA  TCG  CC...
910
...  ARG  GLN  THR  TRP  ARG  SER  LEU  VAL  SER
...  AGA  CAA  ACC  CAG  TGG  CGT  AGT  CTT  GTT  TCT
940
...

```

FIG.6E

```

THR  ASP  GLU  TYR  ALA  THR  LEU  LEU  THR  ASP  ...
ACTGATGAGTATGCAACGCTCTTGACAGAC...
970
...  LYS  ASN  ASN  LYS  PRO  SER  ASP  TYR  ASN  GLY
...  AAAAATAACAAGCCCAAGTGATTACAATGGT
1000
...
ALA  TYR  GLY  HIS  SER  SER  GLU  PHE  ASP  VAL  ...
GCAATAGGTCAATAGCAGTGAAATTGTGATT...
1030
...  ASN  PHE  ALA  ASP  LYS  LYS  ILE  LYS  GLY  LYS
...  AATTTGCTGATAAATAAATAAAGGCAAA
1060
...
LEU  ILE  SER  ASN  GLN  LEU  SER  GLY  THR  ALA  ...
CTTATCAGTAATCAGTTATCAGGCACAGCT...
1090
...  VAL  THR  ALA  LYS  GLU  ARG  TYR  LYS  ILE  GLU
...  GTAACCGCCAAAGAGCGGTATAAATAAGAA
1120
...
ALA  ASP  ILE  HIS  GLY  ASN  ARG  PHE  ARG  GLY  ...
GCTGATATCCACGGCAACCGCTTCCGTGGC...
1150
...  SER  ALA  THR  ALA  SER  ASP  LYS  ALA  GLU  ASP
...  AGTGCCACCGCAAGCGGATAAAGCAGAGAC
1180
...

```

FIG.6F

SER LYS THR GLN HIS PRO PHE THR SER ASP ...
 AGCAAAACCCCAACACCCCTTTACCAAGCAT...
 1210
 ... ALA THR ASN LYS LEU GLU GLY PHE TYR
 ... GCTACAACAAGCTAGAGGTGGTTTATT
 1220
 ... 1240
 ... 1250
 ... 1260

GLY PRO LYS GLY GLU GLU LEU ALA GLY LYS ...
 GGACCAAAAGGCGAGGAGCTGGCAGGTAA...
 1270
 ... PHE LEU THR ASP ASN LYS LEU PHE GLY
 ... TTCTTAACCGATGACACAACAACCTCTTTGGG
 1280
 ... 1300
 ... 1310
 ... 30/73

VAL PHE GLY ALA LYS ARG ASP LYS VAL GLU ...
 GTCTTTGGTGCTAAACGAGATAAGTAGAA...
 1330
 ... LYS THR GLU ALA ILE LEU ASP ALA TYR ALA
 ... AAACCGAAGCCATCTTAGATGCCATGCA
 1340
 ... 1360
 ... 1370
 ... 1380

LEU GLY THR PHE ASN ASN THR ASN LYS ALA ...
 CTGGGACATTTATAATAATACAAATAAGCA...
 1390
 ... THR THR PHE THR PRO PHE THR LYS LYS GLN
 ... ACCACATTCACCCCATTTACCAAAACA
 1400
 ... 1420
 ... 1430
 ... 1440

FIG.6G

LEU ASP ASN PHE GLY ASN ALA LYS LYS LEU ...
 CTGGATAACCTTTGGCAATGCCCAAAGTTTG...
 1450 1460 1470...
 ... VAL LEU GLY SER THR VAL ILE ASN LEU VAL
 ... GTC TTGGGTTCTACCGTCAATTAATTGGTG
 1480 1490 1500
 ...

SER THR ASP ALA THR LYS ASN GLU PHE THR ...
 TCTACCGATGCCCAAAATGAATTCAAC...
 1510 1520 1530...
 ... LYS LYS PHE THR LYS ASP LYS PRO THR SER
 ... AAAAATTCACCAAGACAAAGCCACTCT
 1540 1550 1560
 ...

ALA THR ASN LYS ALA GLY GLU THR LEU MET ...
 GCCACAAACAAGCGGCGGACCTTTGATG...
 1570 1580 1590...
 ... VAL ASN ASP GLU VAL ILE VAL LYS THR TYR
 ... GTGAATGATGAAGTTATCGTCAAAACCTAT
 1600 1610 1620
 ...

GLY LYS ASN PHE GLU TYR LEU LYS PHE GLY ...
 GGCAAAACCTTTGAATACCTAAATTTGGT...
 1630 1640 1650...
 ... GLU LEU SER VAL GLY ASP SER HIS SER VAL
 ... GAGCTTAGTGTCGGTGATAGCCATAGCGTC
 1660 1670 1680
 ...

FIG.6H

PHE LEU GLN GLY GLU ARG THR ALA THR ...
 TTTTACAAAGGCCGAACGCCACCGCTACCAACA...
 1690 1700 1710...
 ... GLY GLU LYS ALA VAL PRO THR GLY LYS
 ... GCGGAGAAAGCCGTAACCAACCAAGCAAA
 1720 1730 1740
 ...
 ALA LYS TYR LEU GLY ASN TRP VAL GLY TYR ...
 GCCAAATATCTGGGGAACCTGGGTAGGATAC...
 1750 1760 1770...
 ... ILE THR GLY ALA GLY THR GLY LYS SER PHE
 ... ATCAGGAGCGGGCAACAGGAAGCTTT
 1780 1790 1800
 ...
 ASN GLU ALA GIN ASP ILE ALA ASP PHE ASP ...
 AATGAGGCCCAAGATATTGCTGATTTTGAC...
 1810 1820 1830...
 ... ILE ASP PHE GLU ARG LYS SER VAL LYS GLY
 ... ATTGACTTTTGAGAGAAATACTAAGGC
 1840 1850 1860
 ...
 LYS LEU THR THR GIN GLY ARG THR ASP PRO ...
 AACTGACCCCAAGGCCGACAGATCCT...
 1870 1880 1890...
 ... VAL PHE ASN ILE LYS GLY GLU ILE ALA GLY
 ... GTCCTTAACATCAAGGTGAATTGCAAGGC
 1900 1910 1920
 ...

32/73

33/73

FIG.6I

```

ASN  GLY  TRP  THR  GLY  LYS  ALA  SER  THR  THR  ...
AATGGCTGGACAGGCAAGCCAGCCACC...
1930
...  LYS  ALA  ASP  ALA  GLY  GLY  TYR  LYS  ILE  ASP
...  AAGCGGACGCAGGAGGCTACAGATAGAT
1940
...  AAGCGGACGCAGGAGGCTACAGATAGAT
1950...
1960
...
1970
1980

```

```

SER  SER  SER  THR  GLY  LYS  SER  ILE  VAL  ILE  ...
TC TAGCAGTACAGGCAATCCATCGTCATC...
1990
...  GLU  ASN  ALA  GLU  VAL  THR  GLY  GLY  PHE  TYR
...  GAATGCCGAGTTACTGGGGCTTTTAT
2000
...  GAATGCCGAGTTACTGGGGCTTTTAT
2010...
2020
...
2030
2040

```

```

GLY  PRO  ASN  ALA  ASN  GLU  MET  GLY  GLY  SER  ...
GGTCCAAATGC AAACGAGATGGGCGGTCA...
2050
...  PHE  THR  HIS  ASP  THR  ASP  SER  LYS  ALA
...  TTACACACGATACCGATGACAGTAAAGCC
2060
...  TTACACACGATACCGATGACAGTAAAGCC
2070...
2080
2090
2100

```

```

SER  VAL  VAL  PHE  GLY  THR  LYS  ARG  GLN  GLN  ...
TCTGTGGTCTTTGGCACAAAGACACAA...
2110
...  GLU  VAL  LYS  ***
...  GAAGTTAAGTAG
2120
...  GAAGTTAAGTAG
2130...
2140

```

10 20 30 40 50 60
MKHIPLTTLCAVLSAVLLTACGSGGS-NPPAPTPIPNASGSENTGN-TGNAGGTNTNAN
.....-.....-.....
.....-.....GSA. A. G-...T...GS.D.
.....S. GF...S...GN. A. -A.....GCANS
.....-.....G. A. A. SG. G. S...A...

	70	80	90	100	
-AGNIGGT-	-----NSGTCGSA				4223
.....S.....S.S.....TP.N.EQ.....				R1
.....K.....DE.K.E.G.....				M5
-V.A.ANSGTCGSAD.EQ.....				LES1
G.A.....	-----CGA.A.S.....K.....DE.K.AE.G.....				Q8
S.A.....	-----S.....S.OK.....D.K.E.G.....				3

110	120	130	140	150
GYGMAISKINLHNRQDTPLDEKNI - ITLDGKKQVA - EGKKSPLPFSLDVENKLLDGVIA				
.....YDQ.....A.....-DNQ.....				
.....M.....KQ.....D.....-K.E.....				
.....A.E.KLR.A.PLNPKNK.AEKR.A.Q.DL.V.DLTNI..DKNLIEY.KKSSEV				
.....VE.KLR.WIPQ-----QEEHAKININDV.KL.DLKHN.DNSTWQNIKNSKEV				
.....VE.KLR.WIPQ-----QEEHAKININDV.KLEGDLKHN.DNSTWQNIKNSKEV				
	160	170		1
	KMWV-ADKNAIGDRIKKG---- <td></td> <td></td> <td></td>			
	...E-.....E...RENEQ..K.			
	...E-.....D---- <td></td> <td></td> <td></td>			
	VSKFE.Q.GG.ENNIRLT---H.DL			
	QTVYNQE.QN.EDQ..RE---- <td></td> <td></td> <td></td>			
	QTVYNQE.QN.ENQ..E-----			

FIG.7B

210	220	230	240	250	
LSSLENKIFHSNDGTTKATTRDLKYVDYGYLANDGNYLTVKTD--KLWNLGPFVGVFY					
...IKA.T....K.....					
.....K.....Q.....V.....--E.....					
QEKYKEL,ENAH,KKSD,RN...E..KS.FNVLSGYTATDHDK---,TNYR,YY,ALY,					
KPTY,KN,NY,H,KQN..R.....RS..IYRSGYSNIIP,-----,IAKT,FD,AL..					
KPIY,KN,ND,H,KQN..R.....RS..IYRSGYSNIDIQK---,IAKT,FD,AL..					
260	270	280	290	300	
NGTTTAKELPTQDAVKYKGHWDFMTDVANRRNRFSEVKENSQA					4223
..S.....KK.....TY...					R1
.....KQ.....L...					M35
K,SE.....QTS,-...Y.....ATLDNKYTDLPGLAR,T					LFS1
Q..Q...Q..VSQ-...T.....AKKGQSFS,FGTSQRL..					Q8
K..Q...Q..VSE-...T.....AKKGQSFS,FERRAGDR					3
					35/73
310	320	330	340	350	
GWYYGASSKDEYNRLLTKEDSAPDGHSGEYGHSSSEFTVNFKEKLTGKLFSLNLDHRH					
..W.....A.A...NY.....E.....S.					
.R.....D.KNK,ERNY.....D.....E.....SR					
Q.-RSLV,T...AT...DKNK,SDYN,A.....D...AD..IK...I..QLSG-					
.DR,S.M,YH..PS...D.KNK,NYN.....D.SK,S,K,E,S..I..G.					
--S.AM..H-..PS...DDKNK,NYND.....D.SK,S...G.....H.					
360	370	380	390	400	
KGNVTKTERYDIDANIHCNFRFRGSATASNKNDTSK--HPFTSDAN					4223
.QK...K...K.D.....D.AED..SK.....K					R1
.....K.....Y.....D.AEA..TK.....K					M35
-TA..AK...K.E.D.....D.AED..TQ.....T					LFS1
..S.N.K.....Y.....DTTEA..SK.....K					Q8
..K...K...N.R.....I..DNE..AK.....D					3

36/73

4223
R1
M35
LES1
Q8
3

FIG.7C

410	420	430	440	450	
NRLEGGFYGPKGEELAGKFLJNDNKLFGVFGAKRESKAEEKTEAILDAYALGTFNISN					
DK.....D.....Q.GNV-					
S.....NA.....G.....KNN					
K.....D.....D-V.....NT.					
S.....NA.....E.K.....KPG					
.....NA.....D.....Q.E.K.-.....KSG					
	460	470	480	490	500
	-ATT--FTPFTKQLDNFGNAKLLVLGSTVIDLVPDTATK--NEFTK----				
	-T.NPA..ANSK.E.....DV.....				
	-...--K.....GV.DV.....				
	K.....K.....N.S.....KFTK				
	-T.NPA..ANSK.E.....G.DV.....				
	-T.NPA..ANSK.E.....IN.....TQG-----D.V.TIDK				

4223
R1
M35
LES1
Q8
3

510	520	530	540	550	
DKPESATNEAGETILMNDVSVKTYG--KNFEYLKFGELSIGGSHSVFLOGERTATTG					
E..K....K.....V.....					
N..D....K.....YGR.....V.T.N.....					
...T...K.....I.....V.D.....					
E..K....K.....I.....YGR.....					
E..ATT..Q...P.T...K.R.QVCC--S.L.H....S...D.N.....K.					
	560	570	580	590	600
	E--KAVPTTGTA [~] KYL [~] GNW [~] GYITGKDTGTGT--GKSFTDAQDVAD [~] DI				
	..--.....K.....A.SSK.STD..G...K.I.....				
	..--.....K.....S-----S...NE.....				
	..--.....K.....A....NE...I.....				
	..--.....E.....S....NE...I.....				
	D.D..M.VA.N...R.T.A..VA.SGNTSKAYEAQQ.A.NANR.E..V				

FIG.7D

610	620	630	640	650	
DFGNKSVSGKLTITKGRQDPV--FSITGQIACNGWIGTASTTKADAGYKIDSSSTGKS					
..EK...N...T.D...--N...E...K...AE.N.....					
..EK...K...T.D...--N...D...K.....					
..ER...K...T.Q.T...--N.K.E...K.....					
..ER...K...T.Q...--N.....A.NV.....					
..A...LT...PNTSS.GKSA.D.AT.D...FS.K.N.PDIET..L...KNSESG					
660	670	680	690	700	
-TVIKDANVTGCFYGPVIANEMGGSFTHNA-----DDSKASVWEGTKRQQEVK-*					
-.....V.....T.....S-----GN.G.V.....K...K*					
-.....E.....T-----T-----E...-*					
-...EN.E.....DT-----DT-----*					
-...EN.K.....DT-----DT-----E...-*					
RVIV...I.I.....Q...L.....YKSNDAIGNQDK...S.....ARK.....P*					

4223
R1
M35
LES1
Q8
3

37/73

FIG.8A

M. catarrhalis strain 4223 *tbpA* - *orf3* - *tbpB* locus gene sequences

```

G A T G C C T G C C T T G T G A T T G G T T G G G G T G T A ...
10                                     20      30...
... T C G G T G T A T C A A A G T G C A A A G C C A A C A G G
40                                     50      60
...

```

tbpA

```

MET ASN GLN SER LYS GLN ASN ...
T G G T C A T T G A T G A A T C A A A C A A A C ...
70                                     80      90...

```

```

... ASN LYS SER LYS LYS SER LYS GLN VAL LEU 38/73
... A C A A A T C C A A A A A T C C A A A C A A G T A T T A 120
...                                     110

```

```

LYS LEU SER ALA LEU SER LEU GLY LEU LEU ...
A A A C T T A G T G C C T T G T C T T T G G G T C T G C T T ...
130                                     140      150...
... ASN ILE THR GLN VAL ALA LEU ALA ASN THR
... A A C A T C A C G C A G G T G G C A C T G G C A A A C A C A 160
...                                     170      180

```

```

THR ALA ASP LYS ALA GLU ALA THR ASP LYS ...
A C G G C C G A T A G G C G G A G G C A C A G A T A G ...
190                                     200      210...
... THR ASN LEU VAL VAL VAL LEU ASP GLU THR
... A C A A A C C T T G T T G T T G T C T T G G A T G A A C T 220
...                                     230      240

```

FIG.8B

```

VAL  VAL  THR  ALA  LYS  LYS  LYS  ASN  ALA  ARG  LYS  ...
G  T  G  T  A  A  C  A  G  C  G  A  A  G  A  A  A  A  C  G  C  C  C  G  T  A  A  A  ...
250                                     260       270...
...  ALA  ASN  GLU  VAL  THR  GLY  LEU  GLY  LYS  VAL
...  G  C  C  A  A  C  G  A  A  G  T  T  A  C  A  G  G  G  C  T  T  G  G  T  A  A  G  G  T  G
280                                     290       300
...

VAL  LYS  THR  ALA  GLU  THR  ILE  ASN  LYS  GLU  ...
G  T  C  A  A  A  C  T  G  C  C  G  A  G  A  C  C  A  T  C  A  A  T  A  A  G  A  A  ...
310                                     320       330...
...  GLN  VAL  LEU  ASN  ILE  ARG  ASP  LEU  THR  ARG
...  C  A  A  G  T  G  C  T  A  A  A  C  A  T  T  C  G  A  G  A  C  T  T  A  A  C  A  C  G  C
340                                     350       360
...

TYR  ASP  PRO  GLY  ILE  ALA  VAL  VAL  GLU  GLN  ...
T  A  T  G  A  C  C  C  T  G  G  C  A  T  T  G  C  T  G  T  G  G  T  T  G  A  G  C  A  A  ...
370                                     380       390...
...  GLY  ARG  GLY  ALA  SER  SER  GLY  TYR  SER  ILE
...  G  G  T  C  G  T  G  G  G  C  A  A  G  C  T  C  A  G  G  C  T  A  T  T  C  T  A  T  T
400                                     410       420
...

ARG  GLY  MET  ASP  LYS  ASN  ARG  VAL  ALA  VAL  ...
C  G  T  G  G  T  A  T  G  G  A  T  A  A  A  A  T  C  G  T  G  T  G  G  C  G  G  T  A  ...
430                                     440       450...
...  LEU  VAL  ASP  GLY  ILE  ASN  GLN  ALA  GLN  HIS
...  T  T  G  G  T  T  G  A  T  G  G  C  A  T  C  A  A  T  C  A  A  G  C  C  A  G  C  A  C
460                                     470       480
...

```

FIG.8C

TYR ALA LEU GLN GLY PRO VAL ALA GLY LYS ...
 1 ATGCCCTACAAAGGCCCTGTGGCAGGCAAA...
 490 500 510...
 ... ASN TYR ALA ALA GLY GLY ALA ILE ASN GLU
 ... AATTATGCCCGCAGGTGGGGCAATCAACGAA
 520 530 540
 ...

ILE GLU TYR GLU ASN VAL ARG SER VAL GLU ...
 ATAGAAATACGAATAATGTCCTCCGCTTGAG...
 550 560 570...
 ... ILE SER LYS GLY ALA ASN SER SER GLU TYR
 ... ATTAGTAAGGTGCAATAATTCAGTGAAATAC
 580 590 600 40/73
 ...

GLY SER GLY ALA LEU SER GLY SER VAL ALA ...
 GGCTCTGGGGCAATTATCTGGGCTCTGTGGCA...
 610 620 630...
 ... PHE VAL THR LYS THR ALA ASP ASP ILE ILE
 ... TTGTTACCAAAACCCGCCGATGACATCATC
 640 650 660
 ...

LYS ASP GLY LYS ASP TRP GLY VAL GLN THR ...
 AAGATGGGTAAAGATTGGGGCGTGCAAC...
 670 680 690...
 ... LYS THR ALA TYR ALA SER LYS ASN ASN ALA
 ... AAACCGCTATGCCAGTAAATAAATACGCA
 700 710 720
 ...

FIG.8D

TRP VAL ASN SER VAL ALA ALA GLY LYS ...
 TGGGTAAATTCTGTGCAAGCAGCAGGCAAG...
 730 740 750...
 ... ALA GLY SER PHE SER GLY LEU ILE ILE TYR
 ... GCAGGTCTTTTAGCGGTCTTATCATCTAC
 760 770 780
 ...
 THR ASP ARG ARG GLY GLN GLU TYR LYS ALA ...
 ACCGACCGCCGTGGTCAAGAAATACAGGCA...
 790 800 810...
 ... HIS ASP ASP ALA TYR GLN GLY SER GLN SER
 ... CATGATGATGCCCTATCAGGGTAGCCAAAGT
 820 830 840
 ...
 PHE ASP ARG ALA VAL ALA THR THR ASP PRO ...
 TTGATAGAGCGGTGGCAACCACTGACCCA...
 850 860 870...
 ... ASN ASN ARG THR PHE LEU ILE ALA ASN GLU
 ... AATAACCGAACATTTTATAAGCAATGAA
 880 890 900
 ...
 CYS ALA ASN GLY ASN TYR GLU ALA CYS ALA ...
 TGTGCCAATGGTAATATGAGCGGTGTGCT...
 910 920 930...
 ... ALA GLY GLY GLN THR LYS LEU GLN ALA LYS
 ... GCTGGCGGTCAACCAACCTTCAAGCCAG
 940 950 960
 ...

41/73

FIG.8E

```

PRO  THR  ASN  VAL  ARG  ASP  LYS  VAL  ASN  VAL  ...
CCA  ACC  AAT  GTG  CGT  GAT  AAG  GTCA  AAT  GTC...
970
...  LYS  ASP  TYR  THR  GLY  PRO  ASN  ARG  LEU  ILE
...  AAG  ATT  ATAC  AGG  TCC  TAA  CCG  CCT  TAT  C
1000
...
PRO  ASN  PRO  LEU  THR  GLN  ASP  SER  LYS  SER  ...
CCA  ACC  CACT  CAC  CCA  GAC  AAT  CCA  AAT  CC...
1030
...  LEU  LEU  LEU  ARG  PRO  GLY  TYR  GLN  LEU  ASN
...  TTA  CTG  CTT  CGC  CCA  GGT  TAT  CAG  CTA  AAC
1040
...
...  LEU  LEU  LEU  ARG  PRO  GLY  TYR  GLN  LEU  ASN
1050
...
...  TTA  CTG  CTT  CGC  CCA  GGT  TAT  CAG  CTA  AAC
1060
...
ASP  LYS  HIS  TYR  VAL  GLY  GLY  TYR  GLU  ...
GAT  AGC  AC  TAT  GT  CGG  TGG  TGT  GAT  GA  A...
1090
...  ILE  THR  LYS  GLN  ASN  TYR  ALA  MET  GLN  ASP
...  ATC  ACC  AA  ACA  AA  A  CTA  CGC  CAT  GCA  AG  AT
1100
...
...  ILE  THR  LYS  GLN  ASN  TYR  ALA  MET  GLN  ASP
1110
...
...  ASP  ILE  GLU  LYS  SER  ARG  LEU  SER  ASN  HIS
1120
...  GAC  ATT  GAA  AA  A  ATCA  AGG  CTA  GCA  ACC  AT
1130
...
...  ASP  ILE  GLU  LYS  SER  ARG  LEU  SER  ASN  HIS
1140
...  GAC  ATT  GAA  AA  A  ATCA  AGG  CTA  GCA  ACC  AT
1150
...
...  ASP  ILE  GLU  LYS  SER  ARG  LEU  SER  ASN  HIS
1160
...  GAC  ATT  GAA  AA  A  ATCA  AGG  CTA  GCA  ACC  AT
1170
...
...  ASP  ILE  GLU  LYS  SER  ARG  LEU  SER  ASN  HIS
1180
...  GAC  ATT  GAA  AA  A  ATCA  AGG  CTA  GCA  ACC  AT
1190
...
...  ASP  ILE  GLU  LYS  SER  ARG  LEU  SER  ASN  HIS
1200
...  GAC  ATT  GAA  AA  A  ATCA  AGG  CTA  GCA  ACC  AT
1210
...

```

42/73

FIG.8F

```

ALA  GLN  ALA  ASN  GLY  TYR  TYR  GLN  GLY  ASN  ...
GCC  AAG  CCA  ATT  GGC  TAT  TAT  CAA  GGC  AAT  ...
1210
...  ASN  LEU  GLY  GLU  ARG  ILE  ARG  ASP  THR  ILE
...  AAT  CTT  GGT  GAA  CGC  ATT  CGT  GAT  ACC  ATT
1220
...
1240
1250
1260

GLY  PRO  ASP  SER  GLY  TYR  GLY  ILE  ASN  TYR  ...
GGC  CAG  ATT  CAG  GTT  ATG  GCA  TCA  ACT  ATT  ...
1270
...  ALA  HIS  GLY  VAL  PHE  TYR  ASP  GLU  LYS  HIS
...  GCT  CAT  GGC  GTA  TTT  TAT  GAT  GAA  AAC  AAC
1280
...
1300
43/73

GLN  LYS  ASP  ARG  LEU  GLY  LEU  TYR  VAL  ...
CAA  AAG  ACC  GCT  AGG  CTT  TGA  ATA  TGT  T...
1330
...  TYR  ASP  SER  LYS  GLY  GLU  ASN  LYS  TRP  PHE
...  TAT  GAC  AGCA  AAG  GTGA  AATA  AAT  GGT  TT
1340
...
1360
1370
1380

ASP  ASP  VAL  ARG  VAL  SER  TYR  ASP  LYS  GLN  ...
GAT  GAT  GTG  CGT  GTG  TCT  TAT  GAT  AAG  CA...
1390
...  ASP  ILE  THR  LEU  ARG  SER  GLN  LEU  THR  ASN
...  GAC  ATT  ACC  GTAC  CGC  AGC  CAG  CAC  CAA  C
1400
...
1420
1430
1440

```

FIG.8G

```

THR  HIS  CYS  SER  THR  TYR  PRO  HIS  ILE  ASP  ...
ACGCACGTGTTCAACCTATCCGGCACATTGAC...
1450
...  LYS  ASN  CYS  THR  PRO  ASP  VAL  ASN  LYS  PRO
...A A A A TGTGTA CGCCCTGATGTC AATAAACCT
1480
...
PHE  SER  VAL  LYS  GLU  VAL  ASP  ASN  ASN  ALA  ...
TTTCGGTTAAAGAGGTGGATACAAATGCC...
1510
...  TYR  LYS  GLU  GIN  HIS  ASN  LEU  ILE  LYS  ALA
...  T A C A A G A C A G C A C A T T T A A T C A A G C C
1540
...
VAL  PHE  ASN  LYS  LYS  MET  ALA  LEU  GLY  SER  ...
GTC TTTAAACA A A A A A TGGCGGTGGG CAGT...
1570
...  THR  HIS  HIS  ILE  ASN  LEU  GIN  VAL  GLY
...  A C G C A T C A T C A C A T C A A C C T G C A A G T T G G C
1600
...
TYR  ASP  LYS  PHE  ASN  SER  SER  LEU  SER  ARG  ...
TAGATAAATTCAATTC AAGCCCTGAGCCGT...
1630
...  VAL  GLU  TYR  ARG  LEU  ALA  THR  HIS  GIN  SER
...  G T A G A A T A T C G T T T G G C A A C C C A T C A G T C T
1660
...

```

44/73

FIG.8H

TYR GLN LYS LEU ASP TYR THR PRO PRO SER ...
 T A T C A A A A C T T G A T T A C A C C C A C C A A G T ...
 1690 1700 1710...
 ... ASN PRO LEU PRO ASP LYS PHE LYS PRO ILE
 ... A A C C C T T T G C C A G A T A A G T T T A A G C C C A T T
 1720 1730 1740
 ...

LEU GLY SER ASN ASN LYS PRO ILE CYS LEU ...
 T T A G G T T C A A A C A A C A A C C C A T T T G C C T T ...
 1750 1760 1770 ...
 ... ASP ALA TYR GLY TYR GLY HIS ASP HIS PRO
 ... G A T G C T T A T G G T T A T G G T C A T G A C C A T C C A
 1780 1790 1800 45/73
 ...

GLN ALA CYS ASN ALA LYS ASN SER THR TYR ...
 C A G G C T T G T A C G C C A A A A C A G C A C T T A T ...
 1810 1820 1830...
 ... GLN ASN PHE ALA ILE LYS LYS GLY ILE GLU
 ... C A A A T T T G C C A T C A A A A A G G C A T A G A G
 1840 1850 1860
 ...

GLN TYR ASN GLN LYS THR ASN THR ASP LYS ...
 C A A T A C A C C A A A A C C C A A T A C C G A T A A G ...
 1870 1880 1890...
 ... ILE ASP TYR GLN ALA ILE ILE ASP GLN TYR
 ... A T T G A T T A T C A A G C C A T C A T T G A C C A A T A T
 1900 1910 1920
 ...

09/673133

FIG.8I

```

ASP  LYS  GLN  ASN  PRO  ASN  SER  THR  LEU  LYS  ...
G A T A A C A A A C C C C A A C A G C A C C C T A A A A ...
1930
...  PRO  PHE  GLU  LYS  ILE  LYS  GLN  SER  LEU  GLY
...  C C C T T T G A G A A A A T C A A C A A A G T T T G G G
1940
...  1960
...  1970
...  1980

```

```

GLN  GLU  LYS  TYR  ASN  LYS  ILE  ASP  GLU  LEU  ...
C A G A A A A T A C A A C A A G A T A G A C G A C T T ...
1990
...  GLY  PHE  LYS  ALA  TYR  LYS  ASP  LEU  ARG  ASN
...  G G C T T T A A A G C T T A T A A A G A T T A C G C A A C
2000
...  2020
...  2030
...  2040

```

46/73

```

GLU  TRP  ALA  GLY  TRP  THR  ASN  ASP  ASN  SER  ...
G A A T G G C G G G T T G G A C T A A T G A C A A C A G C ...
2050
...  GLN  GLN  ASN  ALA  ASN  LYS  GLY  THR  ASP  ASN
...  C A C A A A A T G C C A A T A A A G G C A C G G A T A A T
2060
...  2080
...  2090
...  2100

```

```

ILE  TYR  GLN  PRO  ASN  GLN  ALA  THR  VAL  VAL  ...
A T C T A T C A G C C A A A T C A A G C A A C T G T G T C ...
2110
...  LYS  ASP  ASP  LYS  CYS  LYS  TYR  SER  GLU  THR
...  A A G A T G A C A A A T G T A A A T A T A G C G A G A C C
2120
...  2140
...  2150
...  2160

```

FIG.8J

```

ASN SER TYR ALA ASP CYS SER THR ARG ...
A A G C T A T G C T G A T T G C T C A A C C A C T C G C ...
2170
... HIS ILE SER GLY ASP ASN TYR PHE ILE ALA
... C A C A T C A G T G G T G A T A A T T A T T C A T C G C T
2200
...

```

```

LEU LYS ASP ASN MET THR ILE ASN LYS TYR ...
T T A A A G A C A C A T G A C C A T C A A T A A T A T ...
2230
... VAL ASP LEU GLY LEU GLY ALA ARG TYR ASP
... G T T G A T T T G G G C C T G G G T G C T C G C T A T G A C
2260
...

```

47/73

```

ARG ILE LYS HIS LYS SER ASP VAL PRO LEU ...
A G A T C A A A C A C A A A T C T G A T G T G C C T T T G ...
2290
... VAL ASP ASN SER ALA SER ASN GLN LEU SER
... G T A G A C A C A G T G C C A G C A A C C A G C T G T C T
2320
...

```

```

TRP ASN PHE GLY VAL VAL LYS PRO THR ...
T G G A A T T T G G C G T G G T C G T C A G C C C A C C ...
2350
... ASN TRP LEU ASP ILE ALA TYR ARG SER SER
... A A T T G C C T G G A C A T C G C C T T A T A G A G C T C G
2380
...

```

09/673133

FIG.8K

GLN GLY PHE ARG MET PRO SER PHE SER GLU ...
 C A A G G C T T C G C A T G C C A A G T T T T C T G A A ...
 2410 2420 2430...
 ... MET TYR GLY GLU ARG PHE GLY VAL THR ILE
 ... A T G T A T G G C G A A C G C T T T G G C G T A C C A T C
 ... 2440 2450 2460
 ...
 GLY LYS GLY THR GLN HIS GLY CYS LYS GLY ...
 G G T A A A G G C A C G C A C A T G G C T G T A A G G T ...
 2470 2480 2490...
 ... LEU TYR TYR ILE CYS GLN GLN THR VAL HIS
 ... C T T T A T T A C A T T T G T C A G C A G A C T G T C C A T
 ... 2500 2510 2520
 ...
 GLN THR LYS LYS LEU LYS PRO GLU LYS SER PHE ...
 C A A C C A A G C T A A A A C C T G A A A A T C C T T T ...
 2530 2540 2550...
 ... ASN GLN GLU ILE GLY ALA THR LEU HIS ASN
 ... A A C C A A G A A A T C G G A G C G A C T T A C A T A A C
 ... 2560 2570 2580
 ...
 HIS LEU GLY SER LEU GLU VAL SER TYR PHE ...
 C A C T T A G G C A G T C T T G A G G T T A G T T A T T T ...
 2590 2600 2610...
 ... LYS ASN ARG TYR THR ASP LEU ILE VAL GLY
 ... A A A A A T C G C T A T A C C G A T T T G A T T G T T G G T
 ... 2620 2630 2640
 ...

48/73

FIG.8L

```

LYS  SER  GLU  GLU  ILE  ARG  THR  LEU  THR  GLN  ...
A A A G T G A A G A G A T T A G A A C C C T A A C C C A A ...
2650
...  GLY  ASP  ASN  ALA  GLY  LYS  GLN  ARG  GLY  LYS
...  G G T G A T A A T G C A G G C A A A C A G C G T G G T A A A
2660
...
2680
...

GLY  ASP  LEU  GLY  PHE  HIS  ASN  GLY  GLN  ASP  ...
G G T G A T T G G C C T T T C A T A A T G G A C A A G A T ...
2710
...  ALA  ASP  LEU  THR  GLY  ILE  ASN  ILE  LEU  GLY
...  G C T G A T T T G A C A G G A A T T A A C A T T C T T G G C
2720
...
2740
...

ARG  LEU  ASP  LEU  ASN  ALA  ALA  ASN  SER  ARG  ...
A G A C T T G A C C C T A A A C G C T G C C A A T A G T C G C ...
2770
...  LEU  PRO  TYR  GLY  LEU  TYR  SER  THR  LEU  ALA
...  C T T C C C T A T G G A T T A T A C T C A A C A C T G G C T
2780
...
2800
...

TYR  ASN  LYS  VAL  ASP  VAL  LYS  GLY  LYS  THR  ...
T A T A C A A A G T T G A T G T T A A A G G A A A A C C ...
2830
...  LEU  ASN  PRO  THR  LEU  ALA  GLY  THR  ASN  ILE
...  T T A A C C C A A C T T T G G C A G G A A C A A C A T A
2840
...
2860
...
2880

```

49/73

09/673133

FIG.8M

```

LEU  PHE  ASP  ALA  ILE  GLN  PRO  SER  ARG  TYR  ...
CTGTTTGA TGCCATCCAGCCATCTCGTTAT...
2890
...  VAL  VAL  GLY  LEU  GLY  TYR  ASP  ALA  PRO  SER
...  GTGGTGGGCTTGGCTATGATGCCCAAGC
2920
...
2930
...
2940

GLN  LYS  TRP  GLY  ALA  ASN  ALA  ILE  PHE  THR  ...
CAAAATGGGAGCAACGCCATATTACC...
2950
...  HIS  SER  ASP  ALA  LYS  ASN  PRO  SER  GLU  LEU
...  CATCTGATGCCAAATAATCCAGCGAGCTT
2960
...
2970
...
2980
...
2990
...
3000

LEU  ALA  ASP  LYS  ASN  LEU  GLY  ASN  GLY  ASN  ...
TTGGCAGATAAGAACTTAGGTAATGGCAAC...
3010
...  ILE  GLN  THR  LYS  GLN  ALA  THR  LYS  ALA  LYS
...  ATTCAAACAACAAGCCACCAAGCAAA
3020
...
3030
...
3040
...
3050
...
3060

SER  THR  PRO  TRP  GLN  THR  LEU  ASP  LEU  SER  ...
TCCACGCCGTGGCAACAACCTTGATTGTCA...
3070
...  GLY  TYR  VAL  ASN  ILE  LYS  ASP  ASN  PHE  THR
...  GGTATGTAAACATAAAGATAATTACC
3080
...
3090
...
3100
...
3110
...
3120

```

50/73

FIG.8N

LEU ARG ALA GLY VAL TYR ASN VAL PHE ASN ...
 TTGGCTGGCTGGCTGTAACAATGTTTAAAT...
 3130 3140 3150...
 ... THR TYR TYR THR TRP GLU ALA LEU ARG
 ... ACC TAT TAC ACC ACT TGGGAGGCTTTACGC
 3160 3170 3180
 ...

GLN THR ALA LYS GLY ALA VAL ASN GLN HIS ...
 CAACAGCAAGGGGGCTCAATCAGCAT...
 3190 3200 3210...
 ... THR GLY LEU SER GLN ASP LYS HIS TYR GLY
 ... ACAGGACTGAGCCAGATAGCATTATGGT
 3220 3230 3240
 ...

ARG TYR ALA ALA PRO GLY ARG ASN TYR GLN ...
 CGCTATGCCGCTCCTGGGACGCCAATTACCA...
 3250 3260 3270...
 ... LEU ALA LEU GLU MET LYS PHE ***
 ... TTGGCACTTGAAATGAGTTTAAACCAAGTG
 3280 3290 3300
 ...

GCTTTGATGTGATTTTGGCATGCCAAATCC...
 3310 3320 3330...
 ...CAATCAACCAATGAATAAGCCCCCATTAC
 3340 3350 3360
 ...

51/73

FIG.80

C A T G A G G G C T T A T T T A T C A T C G C T G A G T... 3370
 3380
 ... A T G C T C T T A G C G G T C A T C A C T C A G A T T A G T 3420
 3410
 ...
 C A T T A A T T A T T A G C G A T T A A T T A T T A G T... 3430
 3440
 ... A A T C A C G C T G C T C T T T G A T G A T T T T A A G T G 3480
 3470
 ...
 A T G G G T A T T C A A G A A C G A T G T C A T A C T C A G... 3490
 3500
 ... C A C C G T T T T A T A G G C T T C T A C T T C A A A G A 3540
 3530
 ...
 C A G G C T T G C C T A A A A G T C A T C A A C T T C T A... 3550
 3560
 ... T A T C G C C G A C T T G A T A G C C A C G A G C A G C A A 3600
 3590
 ...
 G C A T T T G A A T G G C T T T T T G A C G A T T T T G G G... 3610
 3620
 ... C A A A G T T G C T G T C G C C A T A A G C T T G T G C T T 3660
 3650
 ...

52/73

FIG.8P

T A A T A C G G T C G T T A G C A A C T G C G G T G G T A G ... 3680
 3670
 ... A G A T A C C A A C G G C A G G C A A C A A A A C A G C A G 3720
 3710
 ...
 C A C T T A G T A C G C C A G C C A A C A G T T T A T T G G ... 3740
 3730
 3750...
 ... T T A A A T T T T C A T A G T A G T T T C C T A A T T A T 3780
 3770
 ...
 T A T C A T T G T A A T T C A T G T T T A T C G T T A T A A ... 3810...
 3800
 3790
 ... A C A A T C G T T A T A A A T A A C T G T G T C G T G A T A 3840
 3830
 ...
 A C C A T T A A T C A C A A G T G G T T A A A T G C C T T ... 3870...
 3860
 3850
 ... T T G C C C A A T G G C A A A T A G G C A C A A T G C T C T 3900
 3890
 ...
 G C T T G T T C T A T G A T G G T C T A T T A T G A T C A T ... 3930...
 3920
 3910
 ... C A T T T T A T T G A C C T A T T T T T A A T C G T A A 3960
 3950
 ...

53/73

FIG. 8Q

TGTTTGTGTTAGTATAAATTTTATC...
3970 3980 3990...
... AATCAACAAATCACAAATATATCAATCAT
4000 4010 4020
AGACGGTAACAGGCTTCATATTTTACGCA...
4030 4040 4050...
... TATTTCCCCAGATGTCGTAGTTTCTATA
4060 4070 4080
GATGATTTGTAAACAAATTGTCGGTCA...
4090 4100 4110...
... TTATCAATTGTAAACTGATGGCTAATTGT
4120 4130 4140
AACCTTATGGCTAATGATAATGAATAA...
4150 4160 4170...
... GCGTTATACCTGTATCAAGAATGAGTAA
4180 4190 4200
ACCATCAATGGTATCTTATTTATCATCAGG...
4210 4220 4230...
... TTGTGTTAATAAGATGCCAATTAGCGACT
4240 4250 4260

FIG.8R

```

A A T T T G T A A A T T A A T A A T C A T T C A T ... 4280
                                     4290...

... A T T T G T A T T T T T A A A T A C C A T A A A A T G G 4320
...                                     4310

orf3
MET LEU ALA PHE LEU ILE GLY ALA ...
T A A A T A T G C T C G C T T T T T G A T A G G A G C T ... 4340
4330                                     4350...

... VAL MET THR ILE THR PRO VAL THR THR 51
... G T C A T G A C A A T C A C G C C T G T T A T A C C A C A 51/73
...                                     4370
                                     4380

PHE THR PRO THR LYS THR PRO ILE LYS PHE ...
T T C A C C C C A C C A A A A C A C C C A T A A A T T T ... 4400
4390                                     4410...
... PHE MET ALA GLY LEU THR PHE LEU ILE ALA
... T T T A T G G C T G G C T T G A C T T T T C T A A T C G C T 4440
...                                     4430
                                     4420

HIS ILE SER HIS ALA ASP ASP GLY ARG THR ...
C A T A T C A G C C A T G C C G A T G A T G T C G C A C C ... 4460
4450                                     4470...
... ASP ASN GLN GLU LEU ILE ASN GLN GLU ILE
... G A C A A T C A A G A G C T A T C A A T C A A G A A T A 4500
...                                     4490
                                     4480

```

FIG.8S

ALA THR LEU GLU PRO ILE ILE ASN HIS ALA ...
 GCCACCCCTTGAAACCCATCATTAACCATGCT...
 4510 4520 4530...
 ... GLN PRO GLU LEU LEU SER HIS ASP ALA LEU
 ... CAGCCCTGAGTTATTGTCTCCCATGATGCATTA
 4540 4550 4560
 ...
 THR PRO LYS ILE GLU PRO ILE LEU ALA GLN ...
 ACCCAAATAAGAACCAATACTGGCAACA...
 4570 4580 4590...
 ... THR PRO ASN PRO ALA GLU ASP THR LEU ILE
 ... ACCCAAATCCCTGCCGAGAGATACGCTCATC
 4600 4610 4620 56/73
 ...
 ALA ASP GLU ALA LEU LEU LEU ASP ASN PRO ...
 GCCGATGAGGCGTTACTGCTTGATTAACCT...
 4630 4640 4650...
 ... ASP LEU LEU ASN HIS ALA LEU ASN SER ALA
 ... GATTGCTCAATCACGCCCTAATAATCTGCT
 4660 4670 4680
 ...
 VAL MET THR ASN HIS MET ALA GLY VAL HIS ...
 GTCATGACCAATCATATGGCAGGCGTTTCA...
 4690 4700 4710...
 ... ALA LEU LEU PRO ILE TYR GLN LYS LEU PRO
 ... GCATTATTGCCCCATTATTCAAAACCTGCC
 4720 4730 4740
 ...

FIG.8T

```

LYS  ASP  HIS  GLN  ASN  GLY  ILE  LEU  LEU  GLY  ...
A A G A C C A T C A A A A T G G C A T T T A C T T G G G ...
4760
...  TYR  ALA  ASN  ALA  LEU  ALA  ALA  LEU  ASP  LYS
...  T A T G C C A A T G C C T T G C C T T T G G A T A A G
4780
...
GLY  ASN  ALA  LYS  LYS  ALA  ILE  ASP  GLU  LEU  ...
G G C A A C G C C A A A A G C C A T T G A T G A G C T A ...
4820
...  ARG  ARG  ILE  ILE  ALA  ILE  MET  PRO  GLU  TYR
...  C G T C G C A T C A T C G C C A T C A T G C C T G A A T A T
4840
...
ASN  VAL  VAL  ARG  PHE  HIS  LEU  ALA  ARG  ALA  ...
A A T G T G G T G C G T T T T C A T C T G G C A A G G C A ...
4880
...  LEU  PHE  MET  ASP  LYS  GLN  ASN  GLU  ALA  ALA
...  T T A T T A T G G A C A A A C A A A A T G A A G C C G C C
4900
...
LEU  ASP  GLN  PHE  ASN  LYS  LEU  HIS  ALA  ASP  ...
C T T G A C C A G T T T A A T A A T T A C A T G C C T G A C ...
4940
...  ASN  LEU  PRO  GLU  VAL  ARG  GLN  VAL  VAL
...  A A C T T G C C A G A G G A G G T G C G G C A G G T T G T T
4980
...

```

57/73

FIG.8U

```

GLY  GLN  TYR  ARG  GLN  ALA  LEU  LYS  GLN  ARG  ...
GGG  CAG  TAC  AGA  CAG  CGCT  A  A  A  C  A  A  C  G  A  ...
4990
...  ASP  SER  TRP  THR  TRP  GLN  VAL  GLY  MET  ASN
...  GAT  TCA  TGG  ACA  TGG  CAA  G  TA  G  GC  A  T  GA  A  T
5000
...  5020
...  5030
...  5040

LEU  ALA  LYS  GLU  ASP  ASN  ILE  ASN  GLN  THR  ...
CTG  GCC  A  A  AG  A  G  A  C  A  A  C  A  T  C  A  A  A  C  C  ...
5050
...  PRO  LYS  ASN  THR  THR  GLN  GLY  GLN  TRP  THR
...  C  C  C  A  A  A  A  C  A  C  C  A  C  G  C  A  A  G  G  T  C  A  A  T  G  G  A  C  T
5060
...  5080
...  5090
...  5100
...  5100 58/73

PHE  ASP  LYS  PRO  ILE  ASP  ALA  ILE  THR  LEU  ...
TTT  GAC  A  A  AAC  C  C  A  T  T  G  A  C  G  C  C  A  T  C  A  C  C  C  T  A  ...
5110
...  SER  TYR  GLN  LEU  GLY  ALA  ASP  LYS  LYS  TRP
...  AGC  TAC  CAA  A  T  T  G  G  G  G  G  C  G  G  A  T  A  A  A  A  G  T  G  G
5120
...  5140
...  5150
...  5160

SER  LEU  PRO  LYS  GLY  ALA  TYR  VAL  GLY  ALA  ...
TCT  TTG  CCC  C  A  A  AG  G  G  C  A  T  A  T  G  T  G  G  A  G  C  G  ...
5170
...  ASN  ALA  GLN  ILE  TYR  GLY  LYS  HIS  HIS  GLN
...  AAC  GCC  C  A  A  A  T  C  T  A  T  G  G  C  C  A  A  A  C  A  T  C  A  T  C  A  A
5180
...  5200
...  5210
...  5220

```

FIG.8V

WO 99/52947

09/673133

PCT/CA99/00307

```

ASN  HIS  LYS  LYS  TYR  ASN  ASP  HIS  TRP  GLY  ...
AATCACAATAAATAACAACGACCATTGGGGC...
5230
...  ARG  LEU  GLY  ALA  ASN  LEU  GLY  PHE  ALA  ASP
...  AGACTGGGGGCAATAATTGGGCTTTGCTGAT
5240
...
5260
...

ALA  LYS  LYS  ASP  LEU  SER  ILE  GLU  THR  TYR  ...
GCCAAAAGAGACCTTAGCATTGAGACCTAT...
5290
...  GLY  GLU  LYS  ARG  PHE  TYR  GLY  HIS  GLU  ARG
...  GGTGAATAAGATTTTATGGGCAATGAGCGT
5300
...
5320
...

TYR  THR  ASP  THR  ILE  GLY  ILE  ARG  MET  SER  ...
TATACCGACACCATTTGGCATATACGCATGTCG...
5350
...  VAL  ASP  TYR  ARG  ILE  ASN  PRO  LYS  PHE  GLN
...  GTGATTATAGAAATCAACCCAAATAATTCAA
5360
...
5380
...

SER  LEU  ASN  ALA  ILE  ASP  ILE  SER  ARG  LEU  ...
AGCCTAAACGCCCATAGACATATCACGCCCTA...
5410
...  THR  ASN  HIS  ARG  THR  PRO  ARG  ALA  ASP  SER
...  ACCAACCATCGGACGCCCTAGGGCTGACAGT
5420
...
5440
...
5460

```

59/73

FIG.8W

```

ASN  ASN  THR  LEU  TYR  SER  THR  SER  LEU  ILE ...
AATAACACTTATACAGTACCTCATTTGATT...
5480
...  TYR  TYR  PRO  ASN  ALA  THR  ARG  TYR  TYR  LEU
...  TATTACCACAATGCCACACGCTATTATCTT  5520
...  5500
...

LEU  GLY  ALA  ASP  PHE  TYR  ASP  GLU  LYS  VAL ...
TTGGGGCAGACTTTTATGATAAAGTG...
5540
...  PRO  GLN  ASP  PRO  SER  ASP  SER  TYR  GLN  ARG  60/73
...  CCAACAAGACCCATCTGACAGTTATCAACGCC  5580
...  5560
...

ARG  GLY  ILE  ARG  THR  ALA  TRP  GLY  GLN  GLU ...
CGTGGCATACGCACAGCGTGGGGCAAGA...
5600
...  TRP  ALA  GLY  LEU  SER  SER  ARG  ALA  GLN
...  TGGCGGGTGCTTTCAGCCGTGCCCAA  5640
...  5620
...

ILE  SER  ILE  ASN  LYS  ARG  HIS  TYR  GLN  GLY ...
ATCAGCATCAACAACAAGCCATTACCAAGG...
5660
...  ALA  ASN  LEU  THR  SER  GLY  GLY  GLN  ILE  ARG
...  GCAACCTAACCAAGCGGTGGACAATAATCGC  5700
...  5680
...

```

FIG.8X

HIS ASP LYS GLN MET GLN ALA SER LEU SER ...
 CATGATAAACAGATGCAAGCGTCTTTATCG...
 5710 5720 5730...
 ... LEU TRP HIS ARG ASP ILE HIS LYS TRP GLY
 ... CTTTGGCACACAGACATTCACCAAATGGGGC
 5740 5750 5760
 ...
 ILE THR PRO ARG LEU THR ILE SER THR ASN ...
 ATCAGCCACGGCTGACCATCAGCAACA...
 5770 5780 5790...
 ... ILE ASN LYS SER ASN ASP ILE LYS ALA ASN
 ... ATCAATAAAGCAATGACATCAGCAAT
 5800 5810 5820
 ...
 TYR HIS LYS ASN GLN MET PHE VAL GLU PHE ...
 TATCACAATAATCAATGTTTGTGTGAGTTT...
 5830 5840 5850...
 ... SER ARG ILE PHE ***
 ... AGTCGCATTTTGTGATGGGATAGCACGCC
 5860 5870 5880
 ...
 CTACTTTTGTTTTGTAAATAATGTGCCA...
 5890 5900 5910...
 ... TCATAGACAATATCAAGAAATAACAAGAA
 5920 5930 5940
 ...

61/73

FIG.8Y

WO 99/52947

09/673133

PCT/CA99/00307

AAAAGATTACAAATTTAATGATAATTGTT... 5950
 ... 5960
 ... 5970...
 ... ATTGTTTATGTTTATTATTATCAATGTAAA 5980
 ... 5990
 ... 6000
 TTGCGGTATTTGTCTATCAATAATGCA... 6010
 ... 6020
 ... 6030...
 ... TTATCAAAATGCTCAATAAATACGCCAAAT 6040
 ... 6050
 ... 6060
 GCACATTGTCAGCATGCCCAAATAAGGCATC... 6070
 ... 6080
 ... 6090...
 ... ACAGACTTTTATGATAATACCATCAACC 6100
 ... 6110
 ... 6120
tbpB
 MET LYS HIS ILE ...
 CATCAGAGGATTATTTTATGAACAACATTC... 6130
 ... 6140
 ... 6150...
 ... PRO LEU THR THR LEU CYS VAL ALA ILE SER A
 ... CTTTAACCACTGTGTGTGCGCAATCTCTG 6160
 ... 6170
 ... 6180

62/73

FIG.8Z

```

LA  VAL  LEU  LEU  THR  ALA  CYS  GLY  GLY  SER  ...
CCGTCCTTATTACCGCTTGTTGGTGGCAGTG...
6190
...GLY  GLY  SER  ASN  PRO  PRO  ALA  PRO  THR  PRO  I
... GTGGTTCAAAATCCACCTGCTCCTACGCCCA 6240
... 6220
... 6230
... 6240

LE  PRO  ASN  ALA  SER  GLY  SER  GLY  ASN  THR  ...
TCCAAATGCTAGCGGTTCAAGGTAACTG...
6250
...GLY  ASN  THR  GLY  ASN  ALA  GLY  GLY  THR  ASP  A 63/73
... GCAACACTGGTTAATGCTGGCGGTA CTGATA 6300
... 6280
... 6290
... 6300

SN  THR  ALA  ASN  ALA  GLY  ASN  THR  GLY  GLY  ...
ATACAGCCAAATGCAGGTAAATACAGCGGTA...
6310
...THR  ASN  SER  GLY  THR  GLY  SER  ALA  ASN  THR  P
... CAACTCTGGTACAGGCA GTGCCAACACAC 6360
... 6340
... 6350
... 6360

RO  GLU  PRO  LYS  TYR  GLN  ASP  VAL  PRO  THR  ...
CAGAGCCAAATAATCAAGATGTACCACCTG...
6370
...GLU  LYS  ASN  GLU  LYS  ASP  LYS  VAL  SER  SER  I
... AGAAAAATGAATAAGATTAAGTTCTCATTCCA 6420
... 6400
... 6410
... 6420

```

FIG.8A'

```

LE  GLN  GLU  PRO  ALA  MET  GLY  TYR  GLY  MET  ...
TTCAAGAACCTGCCATGGGTATGGCATGG...
6430
...ALA  LEU  SER  LYS  ILE  ASN  LEU  HIS  ASN  ARG  G
...CTTTGAGTAAATAATTAATCTACACACCGAC
6460
...
6470
6480

LN  ASP  THR  PRO  LEU  ASP  GLU  LYS  ASN  ILE  ...
AAGACACGCCATTAGATGAATAAATAATCA...
6490
...ILE  THR  LEU  ASP  GLY  LYS  LYS  GLN  VAL  ALA  G
...TTACCTTAGACGGTAATAAACAAGTTGCAG
6500
...
6510
6520
6530
6540
64/73

LU  GLY  LYS  LYS  SER  PRO  LEU  PRO  PHE  SER  ...
AAGGTAAATAATCGCCATTGCCATTTCGT...
6550
...LEU  ASP  VAL  GLU  ASN  LYS  LEU  LEU  ASP  GLY  T
...TAGATGTAGAAATAATAATTGCTTGATGGCT
6560
...
6570
6580
6590
6600

YR  ILE  ALA  LYS  LYS  MET  ASN  VAL  ALA  ASP  LYS  ...
ATATAGCAATAATGAATGTAGCCGATATAA...
6610
...ASN  ALA  ILE  GLY  ASP  ARG  ILE  LYS  LYS  GLY  A
...ATGCCATTGGTGACAGATAATAAGAAAGGTA
6620
...
6630
6640
6650
6660

```


09/673133

FIG.8B'

```

SN  LYS  GLU  ILE  SER  ASP  GLU  GLU  LEU  ALA  ...
A  T  A  A  G  A  A  A  T  C  T  C  C  G  A  T  G  A  A  G  A  C  T  T  G  C  C  A  ...
6670                                     6680
...LYS  GLN  ILE  LYS  GLU  ALA  VAL  ARG  LYS  SER  H
...  A  A  C  A  A  A  T  C  A  A  A  G  A  G  C  T  G  T  G  C  G  T  A  A  A  G  C  C
6700                                     6710
...                                     6720

IS  GLU  PHE  GLN  VAL  LEU  SER  SER  LEU  ...
A  T  G  A  G  T  T  T  C  A  G  C  A  A  G  T  A  T  T  A  T  C  A  T  C  A  C  T  G  G  ...
6730                                     6740
...GLU  ASN  LYS  ILE  PHE  HIS  SER  ASN  ASP  GLY  T
...  A  A  A  C  A  A  A  A  T  T  T  T  C  A  T  T  C  A  A  A  T  G  A  C  G  G  A  A
6750                                     6760
...                                     65/73
...                                     6770
...                                     6780

HR  THR  LYS  ALA  THR  THR  ARG  ASP  LEU  LYS  ...
C  A  C  C  A  A  A  G  C  A  A  C  C  A  C  A  C  G  A  G  A  T  T  A  A  A  T  ...
6790                                     6800
...TYR  VAL  ASP  TYR  GLY  TYR  TYR  LEU  ALA  ASN  A
...  A  T  G  T  T  G  A  T  T  A  T  G  G  T  T  A  C  T  A  C  T  T  G  G  C  G  A  A  T  G
6810                                     6820
...                                     6830
...                                     6840

SP  GLY  ASN  TYR  LEU  THR  VAL  LYS  THR  ASP  ...
A  T  G  G  C  A  A  T  A  T  C  T  A  C  C  G  T  C  A  A  A  C  A  G  A  C  A  ...
6850                                     6860
...LYS  LEU  TRP  ASN  LEU  GLY  PRO  VAL  GLY  GLY  V
...  A  A  C  T  T  T  G  G  A  A  T  T  A  G  G  C  C  C  T  G  T  G  G  T  G  G  T  G
6870                                     6880
...                                     6890
...                                     6900

```

FIG.8C'

AL PHE TYR ASN GLY THR THR ALA LYS ...
 T G T T T A T A T G G C A C A C G A C C G C C A A A G ...
 6910 6920 6930...
 ...GLU LEU PRO THR GLN ASP ALA VAL LYS TYR L
 ... A G T T G C C C A C A C A G A T G C G T C A A A T A T A
 6940 6950 6960
 ...

YS GLY HIS TRP ASP PHE MET THR ASP VAL ...
 A A G G A C A T T G G A C T T A T G A C C G A T G T T G ...
 6970 6980 6990...
 ...ALA ASN ARG ARG ASN ARG PHE SER GLU VAL L
 ... C C A C A G A A G A A C C G A T T T A G C G A A G T G A
 7000 7010 7020 66/73
 ...

YS GLU ASN SER GLN ALA GLY TRP TYR TYR ...
 A A G A A A C T C T C A A G C A G G C T G G T A T T A T G ...
 7030 7040 7050...
 ...GLY ALA SER SER LYS ASP GLU TYR ASN ARG L
 ... G A G C A T C T T C A A A G A T G A A T A C A C C G C T
 7060 7070 7080
 ...

EU LEU THR LYS GLU ASP SER ALA PRO ASP ...
 T A T T A C T A A A G A A G A C T C T G C C C C T G A T G ...
 7090 7100 7110...
 ...GLY HIS SER GLY GLU TYR GLY HIS SER SER G
 ... G T C A T A G C G G T G A A T A T G G C C A T A G C A G T G
 7120 7130 7140
 ...

FIG.8D'

LU PHE THR VAL ASN PHE LYS GLU LYS LYS ...
 A G T T A C T G T T A A T T T A A G G A A A A A A A T ...
 7150 7160 7170...
 ...LEU THR GLY LYS LEU PHE SER ASN LEU GLN A
 ... T A A C A G G T A A G C T G T T T A G T A A C C T A C A A G
 7180 7190 7200
 ...

SP ARG HIS LYS GLY ASN VAL THR LYS THR ...
 A C C G C C A T A A G G C C A A T G T T A C A A A A C C G ...
 7210 7220 7230...
 ...GLU ARG TYR ASP ILE ASP ALA ASN ILE HIS G
 ... A A C G C T A T G A C A T C G A T G C C A A T A T C C A C G
 7240 7250 7260
 ...

67/73

LY ASN ARG PHE ARG GLY SER ALA THR ALA ...
 G C A A C C G C T T C C G T G G C A G T G C C A C C G C A A ...
 7270 7280 7290...
 ...SER ASN LYS ASN ASP THR SER LYS HIS PRO P
 ... G C A A T A A A A T G A C A C A G C A A A C A C C C C T
 7300 7310 7320
 ...

HE THR SER ASP ALA ASN ARG LEU GLU ...
 T T A C C A G T G A T G C C A A C A A T A G G C T A G A A G ...
 7330 7340 7350...
 ...GLY GLY PHE TYR GLY PRO LYS GLY GLU L
 ... G T G G T T T T A T G G G C C A A A A G G C G A G A G C
 7360 7370 7380
 ...

FIG.8E'

```

EU  ALA  GLY  LYS  PHE  LEU  THR  ASN  ASP  ASN  ...
TGGCAGGTAATAATCTTAACCAATGACAAACA...
7390                                7400
...LYS  LEU  PHE  GLY  VAL  PHE  GLY  ALA  LYS  ARG  G
...A  A  C  T  C  T  T  T  G  G  C  G  T  C  T  T  G  G  T  G  C  T  A  A  A  C  G  A  G
7420                                7430
...                                7440

LU  SER  LYS  ALA  GLU  GLU  THR  GLU  ALA  ...
AGAGTAAGCTGAGGAATAAACCGAAGCCA...
7450                                7460
...ILE  LEU  ASP  ALA  TYR  ALA  LEU  GLY  THR  PHE  A
...T  C  T  T  A  G  A  T  G  C  C  T  A  T  G  C  A  C  T  T  G  G  A  C  A  T  T  A
7470                                7480
...                                7490
...                                7500
...                                68/73

SN  THR  SER  ASN  ALA  THR  THR  PHE  THR  PRO  ...
ATACAGTAACGCAACCAATTCACCCCAAT...
7510                                7520
...PHE  THR  GLU  LYS  GLN  LEU  ASP  ASN  PHE  GLY  A
...T  T  A  C  C  G  A  A  A  A  C  A  A  C  T  G  G  A  T  A  A  C  T  T  T  G  G  C  A
7530                                7540
...                                7550
...                                7560

SN  ALA  LYS  LYS  LEU  VAL  LEU  GLY  SER  THR  ...
ATGCCAAATAATGGTCTTAGGTCTTACCG...
7570                                7580
...VAL  ILE  ASP  LEU  VAL  PRO  THR  ASP  ALA  THR  L
...T  C  A  T  T  G  A  T  T  G  G  T  G  C  C  T  A  C  T  G  A  T  G  C  C  A  C  C  A
7590                                7600
...                                7610
...                                7620

```

FIG.8F'

```

YS  ASN  GLU  PHE  THR  LYS  ASP  LYS  PRO  GLU  ...
AAAATGAATTCAACCAAGAGAGT...
7640
...SER  ALA  THR  ASN  GLU  ALA  GLY  THR  LEU  M
...CTGCCAACCAAGAGCGGCGAGACTTTGA
7660
...
7670
7680

ET  VAL  ASN  ASP  GLU  VAL  LYS  THR  ...
TGGTGAATGATGAAGTTAGCGTCAAAACCT...
7690
...TYR  GLY  LYS  ASN  PHE  GLU  TYR  LEU  LYS  PHE  G
...ATGGCAAAACCTTTGAAATACCTAAATTG
7700
...
7710
7720
7730
7740
69/73

LY  GLU  LEU  SER  ILE  GLY  GLY  SER  HIS  SER  ...
GTGAGCTTAGTATCGGTGGTAGCCATAGCG...
7750
...VAL  PHE  LEU  GLN  GLY  GLU  ARG  THR  ALA  THR  T
...TCTTTTACAAGGCGAACGCACTACCA
7760
...
7770
7780
7790
7800

HR  GLY  GLU  LYS  ALA  VAL  PRO  THR  THR  GLY  ...
CAGCGAGAAAGCCGTACCAACCAAGCA...
7810
...THR  ALA  LYS  TYR  LEU  GLY  ASN  TRP  VAL  GLY  T
...CAGCCAATAATTGGGGAACCTGGGTAGGAT
7820
...
7830
7840
7850
7860

```

09/673133

FIG.8G'

```

YR  ILE  THR  GLY  LYS  ASP  THR  GLY  THR  GLY  ...
A  C  A  T  C  A  C  A  G  G  A  A  A  G  G  A  C  A  C  G  G  C  A  ...
                                     7880
...THR  GLY  LYS  SER  PHE  THR  ASP  ALA  GLN  ASP  V
...  C  A  G  G  A  A  A  A  A  G  C  T  T  T  A  C  C  G  A  T  G  C  C  A  A  G  A  T  G
                                     7910
...                                     7900

AL  ALA  ASP  PHE  ASP  ILE  ASP  PHE  GLY  ASN  ...
T  T  G  C  T  G  A  T  T  T  G  A  C  A  T  T  G  A  T  T  T  G  G  A  A  A  T  A  ...
                                     7940
...LYS  SER  VAL  SER  GLY  LYS  LEU  ILE  THR  LYS  G
...  A  A  T  C  A  G  T  C  A  G  C  G  G  T  A  A  C  T  T  A  T  C  A  C  C  A  A  A  G
                                     7970
...                                     7960
                                     70/73

LY  ARG  GIN  ASP  PRO  VAL  PHE  SER  ILE  THR  ...
G  C  C  G  C  C  A  A  G  A  C  C  C  T  G  T  A  T  T  T  A  G  C  A  T  C  A  C  A  G  ...
                                     8010
...GLY  GIN  ILE  ALA  GLY  ASN  GLY  TRP  THR  GLY  T
...  G  T  C  A  A  A  T  C  G  C  A  G  G  C  A  A  T  G  G  C  T  G  G  A  C  A  G  G  G  A
                                     8030
...                                     8020

HR  ALA  SER  THR  THR  LYS  ALA  ASP  ALA  GLY  ...
C  A  G  C  C  A  G  C  A  C  C  A  A  A  A  G  C  G  G  A  C  G  C  A  G  G  A  G  ...
                                     8070
...GLY  TYR  LYS  ILE  ASP  SER  SER  THR  GLY  L
...  G  C  T  A  C  A  A  G  A  T  A  G  A  T  T  C  T  A  G  C  A  G  T  A  C  A  G  G  C  A
                                     8090
...                                     8080

```



FIG.8H'

```
YS  SER  ILE  ALA  ILE  LYS  ASP  ALA  ASN  VAL  ...  
AATCCATCGCCCATCAAGATGCCCAATGTTA...  
8110  
...THR  GLY  GLY  PHE  TYR  GLY  PRO  ASN  ALA  ASN  G  
...  CAGGGGCTTTATGGTCCCAATGCCAATGCG  8160  
...  8140  
  
LU  MET  GLY  GLY  SER  PHE  THR  HIS  ASN  ALA  ...  
AGATGGGCGGCTCATTTACACACACACGCCG...  
8170  
...ASP  ASP  SER  LYS  ALA  SER  VAL  VAL  PHE  GLY  T  71/73  
...  ATGACAGCAAGCCCTCTGTGTCCTTGGCA  8220  
...  8200  
  
HR  LYS  ARG  GLN  GLN  GLU  VAL  LYS  ***  
CAAGAGACAAACAAGAGTTAAGTAGTAAT...  
8230  
...  8240  
...  8250...  
...  TTAACACAAATGTTTG  8260  
...
```

72 /73

FIG.9A

Alignment of *M. catarrhalis* ORF3 proteins

10	20	30	40	50	
MLAFLIGAVMTITPVYTFPTPKTPKFFMAGLTLIAHSHADGRTDN					
.....	60	70	80	90	100
			...P.....G.....T.....		Q8
			QELINQELATLEPIINHAQPELLSHDALTPKIEPILAQTFNPAEDTLIAD		4223
110	120	130	140	150	
EALLDNPDLNHALNSAVMTNMGVHALLPIYQKLPKDHQNGILLGYA					
.....	160	170	180	190	200
			NALAALDKGNAKKAIDELRRIIAIMPEYNNVRFHLARALFMDKQNEAALD		4223
			...V.....A.G.....		Q8
210	220	230	240	250	
QFNKLHADNLPPEVRQVVGQYRQALKQSDSWIWMQVGNLAKEDNINQTPK					
.....	260	270	280	290	300
			NTTQCGWTFDKPIDAITLSYQLGADKKWSLPGAYVGANAQIYGKHHQNH		4223
				Q8
310	320	330	340	350	
KKYNDHMGRLGANLGFADAKKDLSTIETGKRFYCHERYTDTIGIRMSVD					
.....	360	370	380	390	400
			YRINPKFQSLNAIDISRLTNHRTPRADSNNTLYSTSLIYYFNATRYIILG		4223
				Q8

73/73

FIG.9B

410	420	430	440	450	
ADFYDEKVPQDPDSYQRRGIRTAWGQEWAGGLSSRAQISINKRHYQGAN					
.....E.....					
	460	470	480	490	500
	LTSGGQIRHDKQMQASLSLWHRDIHKWGITTPTLTISTNINKSNDIKANH				
Q.....				

4223
Q8

510
KNQMFVFSRIF*
.....*

4223
Q8